Structure-Based Identification of Inhibitors Disrupting the CD2-CD58 Interactions

Tripathi, Neha; Leherte, Laurence; Vercauteren, Daniel; Laurent, Adèle D.

Published in:
Journal of computer-aided molecular design

DOI:
10.1007/s10822-020-00369-z

Publication date:
2021

Document Version
Peer reviewed version

Structure-Based Identification of Inhibitors Disrupting the CD2-CD58 Interactions

Neha Tripathi,[1] Laurence Leherte,[2,3] Daniel P. Vercauteren,[2] Adèle D. Laurent[1*]*

[1]Université de Nantes, CNRS, CEISAM UMR 6230, F-44000 Nantes, France.
[2]Unit of Theoretical and Structural Physical Chemistry, Department of Chemistry, NAmur Research Institute for LIfe Sciences (NARILIS), NAmur MEdicine & Drug Innovation Center (NAMEDIC), Namur Institute of Structured Matter (NISM), University of Namur, Rue de Bruxelles 61, B-5000 Namur (Belgium).
[3]Laboratory of Structural Biological Chemistry, University of Namur, Rue de Bruxelles 61, B-5000 Namur (Belgium).
*Corresponding Authors: Neha Tripathi, Email: Neha.Tripathi@univ-nantes.fr and Adèle D. Laurent, Email: Adele.Laurent@univ-nantes.fr

Abstract

The immune system has very intricate mechanisms of fighting against the invading infections which are accomplished by a sequential event of molecular interactions in the body. One of the crucial phenomena in this process is the recognition of T-cells by the antigen-presenting cells (APCs), which is initiated by the rapid interaction between both cell surface receptors, *i.e.*, CD2 located on T-cells and CD58 located on APCs. Under various pathological conditions, which involve undesired immune response, inhibiting the CD2-CD58 interactions becomes a therapeutically relevant opportunity. Herein we present an extensive work to identify novel inhibiting agents of the CD2-CD58 interactions. Classical molecular dynamics (MD) simulations of the CD2-CD58 complex highlighted a series of crucial CD58 residues responsible for the interactions with CD2. Based on such results, a pharmacophore map, complementary to the CD2-binding site of CD58, was created and employed for virtual screening of ~300,000 available compounds. On the ~6,000 compounds filtered from pharmacophore mapping, ADME screening leads to ~350 molecules. Molecular docking was then performed on these molecules, and fifteen compounds emerged with significant binding energy (< −50 kcal/mol) for CD58. Finally, short MD simulations were performed in triplicate on each complex (i) to provide a microscopic view of the ligand binding and (ii) to rule out possibly weak binders of CD58 from the identified hits. At last, we suggest eight compounds for *in vitro* testing that were identified as promising hits to bind CD58 with a high binding affinity.

Introduction

The immunological response, against pathogenic xenobiotics, graft rejections, or autoimmunity to self-antigens,[1] is a result of a systematic interplay between various immunological cells that include lymphocytes and antigen-presenting cells (APCs). One of the essential phenomena in such processes is the interaction between the T-cells and APCs, followed by the T-cell activation.[2] Their specific adhesion is triggered by the interaction between the cell surface receptors, *i.e.*, lymphocyte function-associated antigen 3 (LFA-3), also named CD58, and the cluster of differentiation 2 (CD2).[3,4] The expression levels of CD2 and CD58 are altered under certain pathological conditions.[5–8] For instance, rheumatoid arthritis patients show an increased expression of CD2 and CD58.[8] Therefore, targeting CD2 and/or CD58 in cell signaling has been identified as an important therapeutic strategy in various fields such as autoimmune diseases,[9] liver transplantation,[10] and cancers.[11] However, differences in their site of expression are observed. The CD2 expression is limited to the lymphocytes T-cells and NK cells while CD58 is widely expressed in various APCs, such as macrophages.[12] Considering its non-hematopoietic expression, targeting CD58 has become the primary approach for the treatment of rheumatoid arthritis and psoriasis.[13] As such, organ transplant rejection has been treated, for instance, with the anti-human CD2 antibodies such as BTI-322 antibody[14] or MEDI-507 antibody.[15]

Considering the wide therapeutic applicability domain, it becomes clear that understanding of the recognition mechanism for CD2 and CD58 could lead to the identification of new disrupting agents. Both macromolecules are transmembrane proteins (Figure 1) and their extracellular domains are involved in the formation of an adhesion complex between T-cells and APCs.[16] A distance of *ca.* 135 Å between the two cells has been mentioned as optimal to allow additional interactions between other receptors of the cell surface for immune-recognition.[16] The binding affinity of CD2 for CD58 is very weak (K_d ~ 1 μM),[16] which facilitates the rapid receptor exchange during intercellular recognition. Although the complete crystal structure of CD2 and CD58 is not available, the interaction domains of the two macromolecules have been experimentally determined (PDB ID: 1QA9).[16,17] The important interaction pairs which help to stabilize the CD2-CD58 complex mainly involved salt bridges formed by residues from each partner, *i.e.*, GluA25-LysB43, LysA32-GluB95, LysA34-AspB32, GluA37-ArgB48, GluA39-LysB51, GluA42-LysB51, ArgA44-AspB31, and GluA78-LysB34 (where chain A is CD58 and chain B is CD2).[17,18] Critical H-bonds between LysA32-AsnB92, AspA33-AsnB92, and AspA84-LysB41 as well

as CH···π interactions between LysA34 and TyrB86 have also been listed as pivotal for the complex formation.[17,18]

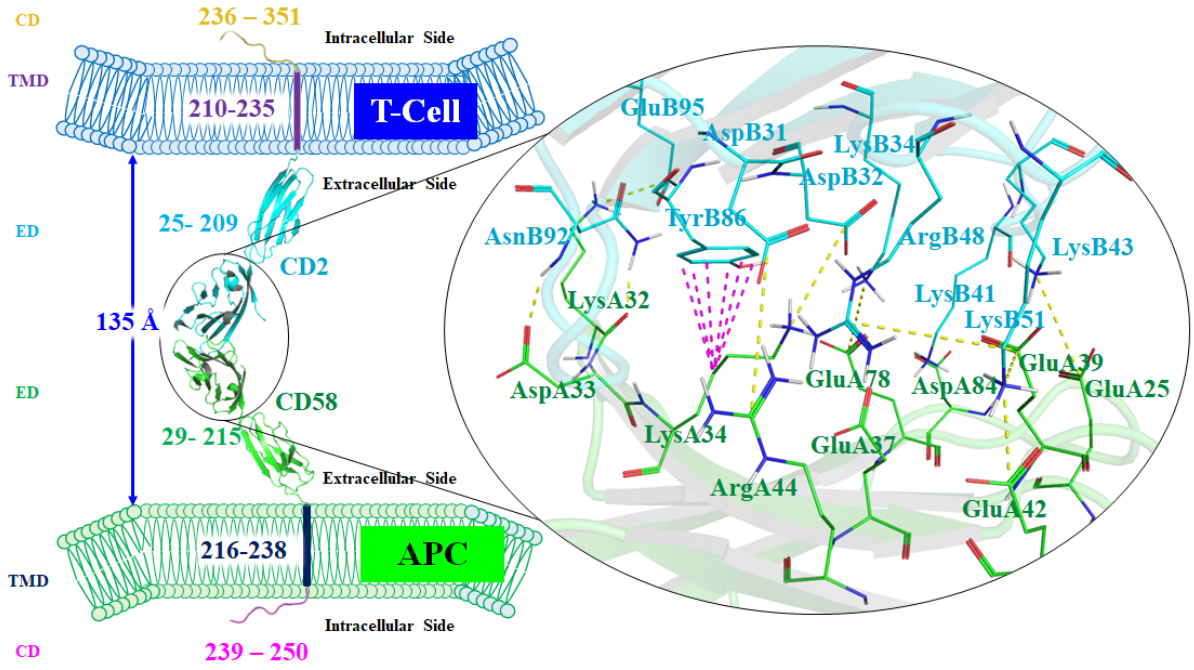

Figure 1. Structural topology and membranous arrangement of CD58 and CD2 (PDB ID: 1QA9). Critical residues involved in the formation of CD2-CD58 complex are shown in the inset as lines, with CD2 in cyan and CD58 in green. H-bonds and CH···π interactions are shown as dashed lines in yellow and magenta, respectively. APC: Antigen-presenting cells, CD: Cytoplasmic domain, TMD: Transmembrane domain, ED: Extracellular domain.

Strikingly, only a few molecular modeling studies have been reported for the structural evaluation of both CD2-CD58 and inhibitor-CD58 complexes. Bayas *et al.* particularly evaluated the force-induced detachment of CD2 and CD58 by employing Steered Molecular Dynamics (SMD) simulations.[18] The authors identified the critical salt bridges for the formation of the CD2-CD58 complex,[17,18] which correlates well with the reported site-directed mutagenesis experiments. Wang *et al.* investigated the effect of CD2-glycosylation on the CD2-CD58 complex formation[19] suggesting that the binding of CD2 to CD58 is governed by the relative position of three loops located at the CD58 binding interface (quantitatively measured as the distance between three atoms, *i.e.*, LysB41-Cα, LysB51-Cα, and GlyB90-Cα in CD2) which in turn is regulated by the glycosylation of AsnB65 (located on the opposite interface). The absence of AsnB65 glycosylation results in an increased distance between these three Cα atoms and therefore, the failure of CD2-CD58 complexation. Considering the tendency to form aggregates and poor solubility of non-glycosylated immunoglobulins, Sun *et al.* proposed apolar-to-polar substitutions in CD58, as a

compensation for the glycosylation without a loss in function for the residues not involved in CD2 binding.[20] Additionally, Lys61Glu, Phe63Leu, and Thr67Ala mutations were incorporated in CD2 to impart structural stability in absence of glycosylation.[17,20] The available crystal structure for CD2-CD58 complex (PDB ID: 1QA9)[17] thus represents the non-glycosylated CD2 and CD58 mutants. Within the drug design framework, several conformationally constrained peptides were already designed to mimic the two CD2 β-strands enriched in charged residues that are interacting with CD58 to avoid the interaction between both partners.[21–25] Their efficiency in inhibiting the CD2-CD58 interaction was evaluated with cell adhesion assays. To understand the atomic-level details of the inhibitor binding with CD58, three known cyclic ligands[21–25] (P6, P7, and RTD-c) mimicking the C and F β-sheets of CD2 were already investigated by our groups, using molecular docking and molecular dynamics (MD) simulations.[26] It was shown that the pivotal residues for the interaction between CD58 and such cyclic inhibitors include AspB32, TyrB86, and AspB87 (from the cyclic inhibitors).[26]

To our knowledge, all efforts made for inhibiting CD2-CD58 interactions were addressed by designing cyclic peptides mimicking the C and F β-sheets of CD2, while no work has been reported yet on the identification of non-cyclic compounds. The present work aims at fulfilling such a gap starting from the available X-ray structures to perform classical MD simulations and revealing key residues contributing to CD2-CD58 stabilization. This step was followed by a structure-based pharmacophore map generation and virtual screening of 227,228 compounds, available from various libraries. After a filtering step using drug-likeness and pharmacokinetic parameters, 343 compounds were identified with satisfactory parameters. A systematic molecular docking was carried out on these compounds, followed by 50 ns MD simulations in triplicate to, at last, clearly identify eight promising compounds disrupting the CD2-CD58 interactions according to our *in silico* strategy.

Computational Methodology

Molecular Dynamics Simulations

The crystallographic structure of the CD2-CD58 complex is available in the RCSB/Protein Data Bank (PDB ID: 1QA9).[17] This structure is a dimer of the CD2-CD58 complex, which contains the adhesion domain from the extracellular segment of CD2 and CD58. The utilized structure is with apolar-to-polar mutations at sites far from the CD2 interacting interface to facilitate the computational experiments with non-glycosylated forms and to maintain the functional state of the CD2-CD58 complex. We have used only one dimer

(CD2-CD58 complex) for the subsequent molecular modeling studies. To evaluate the dynamics of the complex and to identify the residues which are critical for the complex stabilization, classical MD simulations were performed using the AMBER18 package[27] starting from the X-ray structure. Moreover, MD simulations were also carried out on the isolated *apo* forms using the extracted monomers from the CD2-CD58 complex (PDB ID: 1QA9).[17] The various systems were prepared with the help of Amber tools[27] (see Supplementary Information (SI) for details, Table S1). The ionization states of various amino acids, predicted using the *protassign* utility (using PROPKA,[28] pH 7.0 ± 2.0) (discussed in the molecular docking section), were kept during the MD simulations (Figure S1). To maintain uniformity with the previous study by our group,[26] the ligands (obtained after virtual screening) and proteins were treated using the General Amber Force Field (GAFF)[29] and the Amber ff99SB force field,[30] respectively. Additionally, for comparative analysis, GAFF2 (ref. [31]) and FF14SB force field[32] were employed for ligands and protein, respectively, for CD2-CD58 complex and some of the selected hits (discussed later). The terminal residues are located far from the interface of the CD2-CD58 complex and the expected ligand-binding site (Figure S2). The C- and the N-terminal of proteins were deprotonated and fully protonated, respectively. The generated assemblies were submitted to minimization, gradual heating (from 0 to 300 K, under the NVT), and equilibration (1 ns under the NPT, at 310 K and 1 atm pressure). Finally, production runs (500 ns under NPT) were performed for all systems, *i.e.*, the CD2-CD58 complex, and the CD2 and CD58 *apo* forms (pressure relaxation time = 2.0 ps, time step = 2 fs, and cut-off distance for non-bonded interactions = 12 Å). All simulations were run in triplicate to evaluate the reproducibility of the results. The MD results were then analyzed with the help of the *ptraj* module[33] of the Amber[27] and Visual Molecular Dynamics software (VMD).[34] The replicate simulations were submitted to a combined clustering analysis (hierarchical agglomerative (bottom-up) approach, implemented in *ptraj*),[33] in which the root mean square deviation (RMSD) was employed as the clustering parameter. The binding energy (ΔG_{bind}) values were calculated using the Molecular Mechanics-Generalized Born Surface Area (MM/GBSA) method.[35] The complexes between CD58 and hits obtained after virtual screening were also submitted to MD simulations using the same protocol but with a 50 ns production run (top three binding conformations for each ligand, termed as Pose1, Pose2, and Pose3). More details are given in SI.

Pharmacophore Modeling

To identify the crucial structural features and the residues responsible for the formation of the heterodimeric complex between CD2 and CD58, pharmacophore mapping was employed using the crystal structure of the CD2-CD58 complex. The surface area between the two macromolecules is *ca.* 1,200 Å^2, with roughly equal contribution from both the partners.[17] Arulanandam *et al.* reports a surface area of 770 Å^2 for the CD58-binding site.[36] To make this step efficient, pharmacophoric feature selection was guided by the binding energy and per-residue decomposition energy calculations performed following the MD simulations. The CD58 residues which contributed to the binding with CD2 by an energy equivalent to or less than –1.0 kcal/mol (in all the three replicate runs) and participated in close interaction with CD2 surface were considered to be important for the development of the pharmacophore map. Selected residues are LysA32, AspA33, ArgA44, and PheA46. The pharmacophore model was then defined based on the selected nine potential pharmacophore features (discussed later) complementary to the CD2-binding site using the *Phase* module[37,38] of the Schrödinger software package.[39]

The 3D structures of the 227,228 compounds, collected from various small molecular libraries (Table S2) were prepared using the LigPrep module of Schrödinger.[40,41] In such a process, their ionization states were generated using Epik ionizer[42–44] (at pH 7.0 ± 2.0). Subsequently, the prepared ligands were submitted for the prediction of conformers using the ConfGen utility of Schrödinger,[45,46] the maximum number of conformers to be generated being kept to 10. For the pharmacophore hypothesis based screening, functionally similar features, *i.e.*, (i) hydrophobic groups and aromatic rings, (ii) H-bond donating (HBD) and positive functional groups, were treated equivalently. From the 227,228 compounds considered in the ligand dataset, 5,882 unique molecules simultaneously exhibited at least six out of the selected nine features, defined by the pharmacophore model, and were then further taken for the next step of virtual screening.

Property-Based Filtering to Identify Drug like Compounds

The molecular descriptors of the small molecules can be correlated to their physicochemical properties, which in turn govern their pharmacokinetic behavior. Considering this, property-based filtering was employed regarding the molecules selected after the pharmacophore map-based screening. In the first step, the Drug Likeness Tool (DruLiTo)[47] was utilized. The filtered compounds were submitted to the calculation of various molecular descriptors and subsequent filtering based on various drug-likeness rules

including the Lipinski's rule,[48] Ghose filter,[49] Veber rule,[50] MDDR like rule,[51] BBB likeness,[52] unweighted quantitative estimation of drug likeness rule (uwQED), and weighted QED (wQED).[53] Compounds with uwQED > 0.4 were taken for the next step of screening, in which the *QikProp* module[54] of Schrödinger software was employed to calculate the pharmaceutically important properties of small molecules (Table S3). The selection criteria for the next step of virtual screening, *i.e.,* molecular docking, was set as #metabol <5, percent absorption >80%, and QpLogHERG > −5.

Molecular Docking Based Virtual Screening

In a third step, a systematic molecular docking based screening protocol was adopted. To begin with, the pre-processing of the CD58 macromolecular structure (last frame from the complex after MD simulations, which belongs to the most populated cluster in the MD simulation trajectory, as discussed later) was performed, using the Protein Preparation Wizard module[40] of Schrödinger,[55] in which missing hydrogens were added, water molecules beyond 5 Å were removed, and the correct bond orders were assigned. This step was adopted to parameterize the atoms for molecular docking. The protonation state and orientation of the side chain functional groups of amino acids were predicted with the help of the *protassign* utility (using PROPKA,[28] pH 7.0 ± 2.0), which assigns the same p*Ka* values as predicted by the PDB2PQR server (http://server.poissonboltzmann.org/pdb2pqr).[56] This was followed by a restrained minimization (cut-off RMSD = 0.3 Å) with the *impref* utility. The ligands obtained after the property-based screening were submitted to a structure preparation using the LigPrep module of Schrödinger.[40,41] In such a process, their protonation states were generated using Epik ionizer [42–44] (at pH 7.0 ± 2.0).

The interaction site for the small molecules was defined at the centroid of the above-mentioned residues, as identified from the MM/GBSA ΔG_{bind} and per-residue decomposition energy calculations (grid center: 6.42, 54.59, 56.35; inner box: 10 $Å^3$, outer box: 20 $Å^3$) using the Receptor Grid Generation module of Schrödinger. *Glide* of Schrödinger[57–60] was employed for the molecular docking based virtual screening performed in three sequential steps, *i.e.*, (i) high-throughput virtual screening (HTVS) (one pose per ligand), (ii) Standard Precision (SP) mode docking (20 poses per ligand), (iii) eXtra Precision (XP) mode docking (20 poses per ligand), and (iv) Induced Fit Docking (IFD) (20 poses per ligand).[61–63] All parameters were set to default values and the OPLS_2005 force field[64] was used for the molecular docking stage. Prime MM/GBSA calculations for the binding energy were

performed after the SP mode docking and the molecules with ΔG_{bind} values < −35 kcal/mol were submitted to the XP mode semi-flexible molecular docking, for which the selection criteria was set to Prime MM/GBSA ΔG_{bind} values < −40 kcal/mol. The pose filtering was guided by the presence of the crucial interactions with CD58 and reproducibility of the docked conformations. This resulted in the selection of fifteen molecules for the next step of virtual screening. The selected fifteen molecules, which will be referred to hereafter as **compounds 1** to **15**, were submitted to flexible molecular docking (IFD) (grid at the centroid of LysA32, AspA33, ArgA44, and PheA46; inner box: 10 $Å^3$, outer box: 30 $Å^3$; distance cut-off: 5 Å; molecular docking in SP mode), for which the Prime MM/GBSA ΔG_{bind} value cut-off was set to < −50 kcal/mol.

Results and Discussion

Key residues stabilizing CD2-CD58 interactions

The crystallographic structure for apo-CD2, apo-CD58, and the CD2-CD58 complex are available in RCSB/PDB (Table S4). Interestingly, the complexed forms of CD2 and CD58 have a few mutated residues as compared to the apo-forms of CD2 and CD58 (Figure S3), to impart structural stability in the absence of glycosylation.[17,20] In the CD2-CD58 system, the three residue mutations, incorporated in CD2 are Lys61Glu, Phe63Leu, and Thr67Ala, whereas the incorporated mutations in CD58 are Val21Gln, Thr63Ser, Met65Ser, and Leu93Gly. Figure 2 depicts the structural superimposition of *apo*-CD58 (in magenta) and CD2-bound CD58 (in cyan) which helps to identify the residues exhibiting a conformational change between the unbound and bound CD58. Looking at the binding CD2-CD58 interface, several residues such as Phe1 (Ser1 in *apo*-CD58), Asn20, ValA21 (GlnA21 in *apo*-CD58), GluA25, GluA39, AsnA40, ArgA44, GluA76, GluA78, AsnA81, and AspA84 change their conformations in the presence/absence of CD2. Among them, GluA25, GluA39, and GluA78 are known to be involved in the formation of salt bridges in the CD2-CD58 system (Figure 1).[17,18]

The resolution of the available CD2-CD58 crystal structure is 3.2 Å (Table S4). There are 208 close contacts (with clash overlap > 0.4 Å) in the complete structure, with 76 intramolecular close contacts (43 in CD58 and 33 in CD2) and 9 intermolecular close contacts, within the dimer used for the molecular modeling studies in this work (Table S5).[17] Therefore, a 500 ns MD simulation (in triplicate) was performed starting from the crystallographic structure of the CD2-CD58 complex (PDB ID: 1QA9) to resolve the close

contacts in the structure and to evaluate the key features of the CD2-CD58 binding. In our previous work,[26] the cyclic peptides were observed to shift from the initial and expected site of binding in CD58. Considering this, an additional purpose of performing MD simulations of the CD2-CD58 complex was to identify the most stable intermolecular interaction during the simulations and select the residues with a higher contribution in terms of binding energy in the formation of the complex.

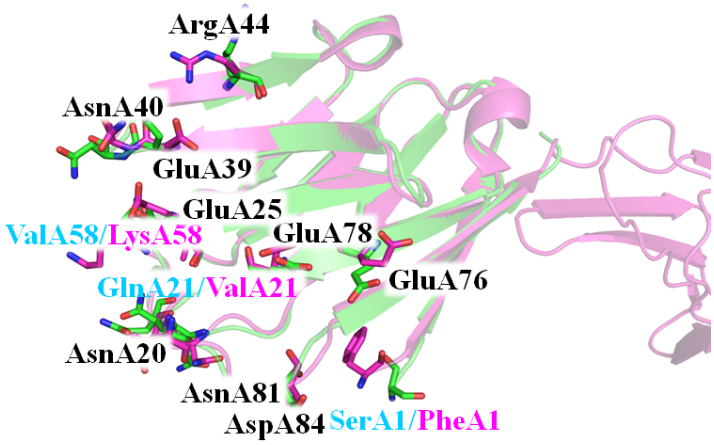

Figure 2. Structural superimposition of *apo*-CD58 (magenta, PDB ID: 1CCZ) and CD2-bound CD58 (green, PDB ID: 1QA9). The residues with differing conformations are shown as sticks. The CD58 residues which are mutated in the CD2-CD58 complex are labeled in magenta and cyan, for the *apo*-CD58 and CD2-bound CD58.

Combined clustering analysis for the CD2-CD58 complex (Figure S4A) shows that the results for two replicates (named Run1 and Run2) are highly similar, as indicated by the presence of more than 90% of the frames in one cluster (C#0). The third replicate (Run3) exhibited a distribution of structures spanned over two clusters (48.14% of C#0 and 51.5% of C#1) (Figure S4A). To identify the component responsible for non-reproducible results during the MD simulations, an additional clustering analysis was performed based on CD2 and CD58, separately. CD2 exhibits a distribution of coordinates over a single cluster for all the replicated simulations (Figure S4B) while CD58 exhibited a wider distribution of structure over multiple clusters (Figure S4C). The RMSD value (Table S6) between the centroid frame of clusters C#0 and C#1 (in the combined cluster analysis of three replicate simulations) was 1.74 Å (calculated for CD2 and CD58) when aligned on the whole CD2-CD58 complex. On the other hand, the alignment on CD58 or CD2 showed a RMSD value of 1.15 Å and 1.88 Å (Table S6), respectively. The system stability analysis ensured that the CD2-CD58 complex was well equilibrated during the simulation period for all replicates (Figure S5). The whole system RMSD for the complex was less than 3.5 Å over the

simulation period, as compared to the initial structure. The analysis of the backbone RMSD for CD2 and CD58 in the complex showed an RMSD below 2 Å, indicating that the fluctuations are mainly arising due to side-chain flexibility of the surface residues. The insignificant difference observed in the radius of gyration of the complexes ensures a stable complex formation during the MD. The distance between the center of mass (COM) for the CD2 and CD58 in the complex was nearly constant in all three replicate simulations, indicating a stable complex formation. It is observed that the absence of glycosylation does not affect the formation of the CD2-CD58 complex, when the apolar-to-polar mutations,[20] are incorporated in the macromolecules according to the in vitro experiments. The computed B-factor and the atomic fluctuations show that CD2 exhibited much lower dynamics as compared to CD58, which is in agreement with the experimental B-factors. To understand the per-residue fluctuation pattern of the two macromolecules, the *apo*-CD2 and *apo*-CD58 structures were also submitted to MD for 500 ns, in triplicate. The reproducibility and system stability analyses (Figure S6) confirmed that the three simulations behaved similarly for *apo*-CD2 and *apo*-CD58. A comparative analysis of atomic fluctuation (Figure S7) in *apo*-CD2, *apo*-CD58, and CD2-CD58 complex showed that CD2 exhibits a similar atomic fluctuation. On the other hand, CD58 in the complex exhibits a much higher fluctuation as compared to *apo*-CD58 (Figure S7A-S7C). The analysis of the location for the residues with atomic fluctuation > 3 Å (Figure S7D) showed that most of these residues are located at the protein surface (Figure S7E), and thus have more flexibility as they are exposed to the solvent in the simulation.

After ensuring that the MD simulation trajectories for the CD2-CD58 complex have reached stability along the simulation run, we proceeded to an in-depth analysis of the interactions between the two partners. The binding energy analysis performed for each of the 20 ns indicates a highly fluctuating pattern (Figure S8). The MM/GBSA ΔG_{bind} values for Run1 and Run2 were similar to each other but not with Run3. However, the ΔG_{bind} values between 400 and 500 ns were observed to be stable for the three replicates. The average ΔG_{bind} for the CD2-CD58 complexation over the last 100 ns for Run1, Run2, and Run3 are –22, –22, and –42 kcal/mol, respectively, whereas that for the most populated clusters (Figure S9) identified during the individual clustering analysis (based on interfacial residues from CD2 and CD58) was calculated to be –23, –18, –44 kcal/mol, for Run1, Run2, and Run3, respectively. The residue-wise decomposition energy values were analyzed to identify the pivotal residues with a reproducible significant contribution in the CD2-CD58 complex. This

was achieved (i) through the occupancy analysis (frequency of residue occurrence during the various calculations) for the various residues in three simulations (Figure 3A), and (ii) by evaluating the averaged per residue decomposition energy values for the most populated cluster (C#0 for Run1, Run2, and Run3 in Figure S9).

Regarding the occupancy analysis, the decomposition energy for each residue was averaged over every 20 ns (counted as one calculation). The residues which exhibited binding energy < 1.0 kcal/mol (with CD2) for more than 30% of the calculations were considered as important for the CD2-CD58 complex formation. In this process, seven CD58 residues possess a higher occupancy in all three replicate simulations: LysA32, AspA33, GluA37, GluA39, ArgA44, PheA46, and ProA80 (Figure 3A). Among them, two have never been reported as playing a key role in the CD2-CD58 interactions, *i.e.*, PheA46 and ProA80. Based on such analysis, GluA25, LysA29, GluA76, ThrA83, and AspA84, showed a significant contribution to the CD2-CD58 complex formation during Run3, but not in Run1 and Run2, and hence could be responsible for the observed lower binding energy of CD2 for CD58 during Run3. The per-residue decomposition energy analysis showed that the contribution from these residues (except for GluA37 and GluA39) was considerably more stable throughout the simulations for the three replicates (Figure S10).

Similarly, seven CD2 residues that were involved in the CD58 interaction were identified from the occupancy analysis for the three replicates (Figure S11): AspB31, AspB32, ArgB48, TyrB86, GlyB90, LysB91, and AsnB92. Most of these CD2 residues were involved in salt bridges or H-bonds in the crystallographic structure while GlyB90 backbone (H-bond with LysA34 of CD58) and AsnB92 side chain (H-bond with AspA32 of CD58) were detected as the most stable interactions *via* Steered MD.[18] Employing the second method, *i.e.*, the average per residue decomposition energy analysis for the most populated clusters for the three replicates resulted into the identification of LysA32, GluA37, GluA39, ArgA44, PheA46, AspB31, AspB32, TyrB86, GlyB90, LysB91, and AsnB92 (Figure S12) as the important residues for the complex formation, which are overlapping with the ones identified from the first method in this work. To understand the observed differences between Run3 and the other two replicates, we also performed a clustering of the simulation trajectories based on the all-atom RMSD of the interfacial residues selected after decomposition energy analysis. Out of the simulations, Run1 and Run2 show similar results during the combined clustering analysis, whereas Run3 shows a significant difference in the configuration of the interfacial residue from CD2 over the last 250 ns, where the CD2

conformation spanned over two clusters (Figure S13). Such a distribution was not observed in Run1 and Run2. In the combined clustering, the frames extracted from Run3, equally span over C#0 (47%) and C#1 (53 %) for CD2, whereas for Run1 and Run2, the C#0 is the majorly populated cluster (81% and 100%, respectively) (Figure S13). It is very likely the origin of the significant differences in terms of binding energy.

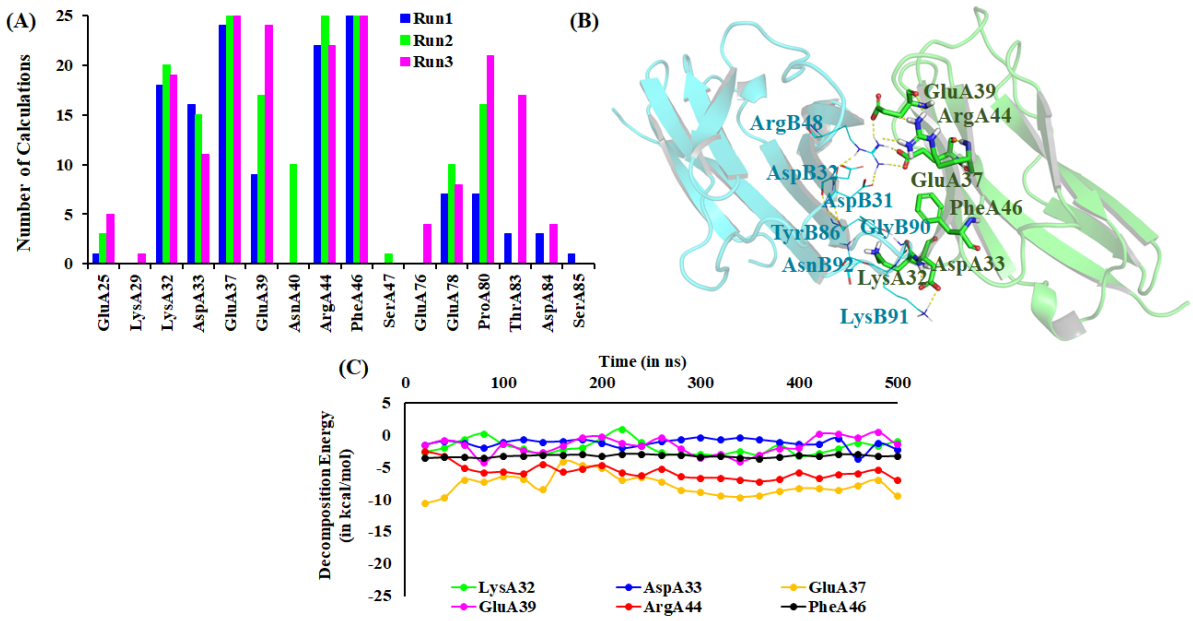

Figure 3. Identified key residues in the CD2-CD58 interaction (A) Occupancy analysis for the CD58 residues contributing to the ΔG_{bind} value for CD2 by more than −1.00 kcal/mol in the three simulations (Run1, Run2, and Run3); (B) selected CD58 residues (shown as sticks) and the important residues from CD2 (shown as lines); and (C) decomposition energy for the selected CD58 residues over the period of 500 ns MD simulation for replicate Run2.

The topographical positions for the identified residues are shown (as lines) in Figure 3B. It can be clearly stated that although ProA80 from CD58 was identified to have a significant contribution to the decomposition energy (Figure 3A), the selected CD2 residues were not involved in direct interaction with ProA80 (Figure 3B). The observed low energy is due to the intramolecular interactions of ProA80 within CD58 and was therefore not considered for the analysis. The per-residue decomposition energy for the CD58 residues identified from the occupancy analysis, in the three replicate simulations shows that Run1 and Run3 exhibited a higher degree of fluctuation in the energy contribution of GluA37 and GluA39 (Figure S10). On the other hand, Run2 exhibited a stable energy contribution for all six residues after 300 ns of the simulation period (Figure 3C). Therefore, the observed order in the binding affinity during three simulation replicates is Run3 > Run1 > Run2. Considering the variation in the energy contribution between the replicate runs, both residues were not

considered later for the pharmacophore development. Therefore, solely four CD58 residues were taken into account for the development of the pharmacophore map, *i.e.*, LysA32, AspA33, ArgA44, and PheA46.

Pharmacophore Mapping and Virtual Screening

As discussed in the previous sections, only a part of the entire interface in the CD2-CD58 complex is responsible for the binding of the two macromolecules (Figure 3B). Therefore, the pharmacophore map generation was performed at the centroid of the key CD58 residues LysA32, AspA33, ArgA44, and PheA46 (Figure 4A). A total of fifteen pharmacophoric features were detected: four hydrogen bond acceptors (HBA, noted by A_x), five hydrogen bond donors (HBD, noted by D_x), two hydrophobic functional groups (noted by H_x), one negatively charged group (noted by N_x), two positively charged groups (indicated by P_x) and one aromatic feature (indicated by R_x). From these fifteen features, nine of them were selected regarding their relative positions on the CD58 surface, as well as their relative position to the others. The selected features include two HBA (A_1 and A_4), three HBD (D_6, D_7, and D_9), two hydrophobic groups (H_{10} and H_{11}), one negative functional group (N_{12}), and one aromatic group (R_{15}) (Figure 4B). The selection was optimized to keep the generated pharmacophore map compact and include the non-overlapping features only. From these features, A_1 and H_{11} corresponded to LysA32, D_6, and N_{12}, to AspA33, D_7, and H_{10}, to the side chain guanidine group of ArgA44, and A_4, D_9, and R_{15} belonged to PheA46 (Figure 4A). Since the ligands interacting at a particular binding site are assumed to exhibit complementary features to the HBD, HBA, and ionic functional groups, the generated map was manually modified to reverse the interacting nature of these groups. The finally generated hypothesis is shown in Figure 4B, in which the HBAs are replaced by HBDs, HBDs by HBAs, and negative functional groups are exchanged with positive ones. The generated complementary pharmacophore hypothesis was employed for the screening of the compound library (Table S2).

To identify new CD58-binding agents that might specifically inhibit the CD2-CD58 interactions, a systematic virtual screening protocol (Figure S14) was adopted (see computational methodology section for details). In this process, the first step was the pharmacophore map-based screening to identify the molecules which exhibited complementary features to that of the CD2-binding site of CD58. We utilized the last structure from the MD simulations, which belonged to the most populated cluster (Figure S4), and exhibited an RMSD of 1.43 Å with the centroid of this cluster (Figure S15). There

were no significant differences at the CD2 interacting interface, especially regarding the residues considered as important for virtual screening protocol. The presence of six features out of the selected nine features was considered as the criterion for the pharmacophore-based filtering. A total of 5,882 molecules out of 227,228 compounds exhibited a minimum of the selected six features. The structural superimposition of the top twenty molecules (based on the PhaseScreen score) with the generated pharmacophore map is presented in Figure S16. Features H_{11}, D_{c1}, A_{c6}, and R_{15} were present in almost all the selected compounds. An interesting categorization of molecules was observed among the filtered molecules, which (i) either exhibited H_{11}, D_{c1}, A_{c6}, R_{15}, D_{c4}, and A_{c9} (green in Figure 4D) (ii) or exhibited H_{11}, D_{c1}, A_{c6}, R_{15}, H_{10}, and A_{c7} (magenta in Figure 4D).

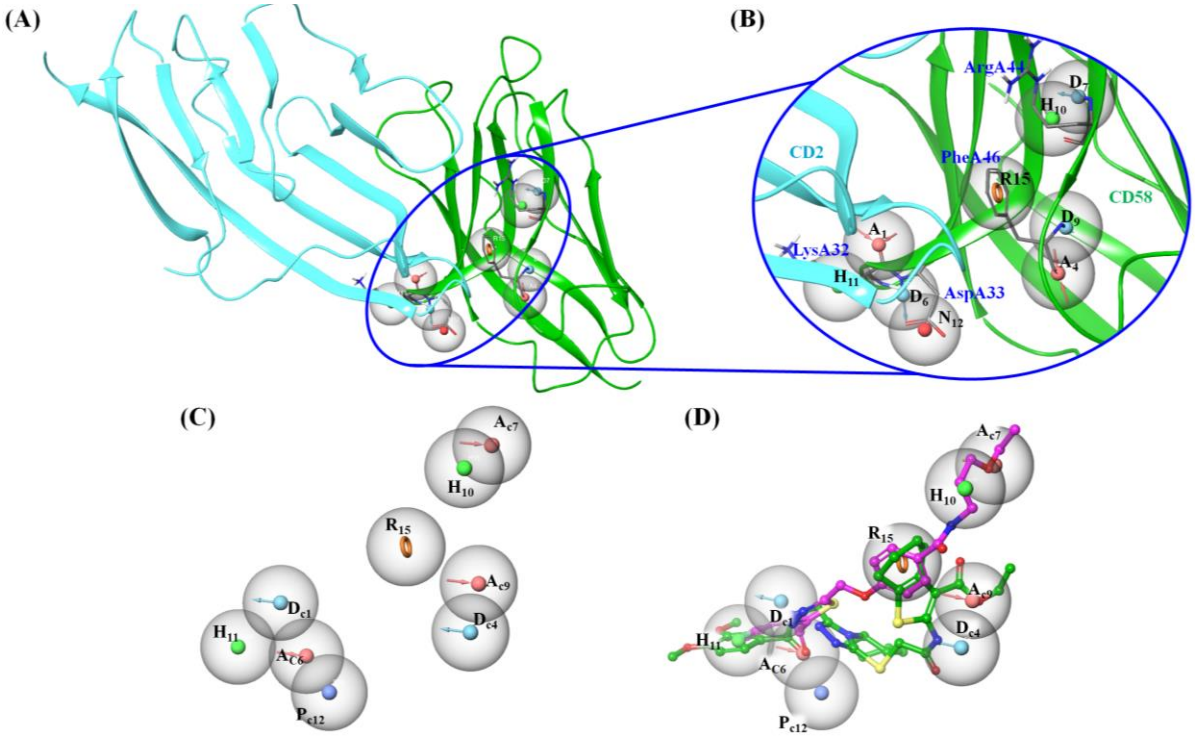

Figure 4. Pharmacophore map-based virtual screening to identify the CD58 binding moieties. (A) Identified structural features, which are responsible for binding CD2 (cyan), at the CD58 surface (green). Residues considered for the development of the pharmacophoric hypothesis shown as sticks. (B) The pharmacophoric hypothesis, which was considered for virtual screening, features complementarity to the CD2-binding site in CD58. (C) Two molecules (in magenta and green) aligned to the generated pharmacophore to represent the molecular features that are majorly present in the selected compounds.

The filtered 5,882 molecules were then submitted for the prediction of their physicochemical characteristics and thus their pharmacokinetic properties. Such a principle was the basis of many drug-likeness evaluation tools, which evaluate the suitability of a particular compound according to reported filters. One interesting tool is DruLiTo[47] which

along with the several well-known drug-likeness filters also incorporates the quantitative estimation of drug-likeness in the filtering process (see SI for the property-based filtering details, Table S7, Figure S17-S18). After calculating the molecular properties of the various molecules, 343 unique compounds were identified to be suitable as plausible drug candidates and were taken for the next step of virtual screening, *i.e.*, molecular docking at the CD2-binding site in CD58.

Molecular Recognition of the Identified Hits after Molecular Docking

To evaluate the molecular recognition potential of the selected compounds for CD58, molecular docking studies were performed with an increasing degree of precision. Out of the 343 molecules submitted to the HTVS procedure, 269 exhibited a docking score < −2, which were taken for semi-flexible docking. During the Standard Precision (SP) mode docking, 41 molecules exhibited a reproducible binding conformation and a Prime MM/GBSA ΔG_{bind} values < −35 kcal/mol, which were then subjected to the eXtra Precision (XP) mode. From these 41 molecules, fifteen exhibited a Prime MM/GBSA ΔG_{bind} values < −50 kcal/mol. The 2D structures of these fifteen compounds are presented in Figure S19. Considering the non-flexibility of the protein side chains in SP and XP mode, the filtered fifteen molecules were submitted to the Induced Fit Docking (IFD) (with SP mode) to take the amino acid side chain flexibility into account.

Molecular docking showed that the fifteen molecules occupied the CD2-binding interface in a wide range of orientation (Figure S20). To accommodate the docked molecules, the loops present on the CD2-binding surface of CD58 were considered to be flexible (green in Figure S20). The results for the docking using IFD (for the pose with the lowest value of Prime-MM/GBSA ΔG_{bind}) are presented in Table 1. The 3D interaction diagrams for **compounds 1-15** are shown in Figure 5. 2D-interaction diagrams are also presented in Figures S21-S22. The top-scoring **compound 1** exhibited a Prime-MM/GBSA ΔG_{bind} of −65 kcal/mol and a glide gscore of −7.91. The molecular recognition interactions which stabilized the complex of **compound 1** with CD58 include H-bond interactions with LysA34, SerA47, SerA48, and NH···π interactions with LysA30. Other residues which participated in the accommodation of **compound 1** in the CD2-binding site are LysA32, AspA33, ValA35, PheA46, PheA49, SerA70, AspA71, and TyrA75. For most of the selected molecules, LysA30, LysA32, AspA33, LysA34, and SerA47 from CD58 participated in H-bonds/salt bridges/CH···π/ NH···π interactions. The selected compounds were rich in H-bond accepting

groups (Figure S19), which facilitated their accommodation and interactions at the CD58 surface.

Table 1. Molecular docking results from the IFD procedure of **compounds 1-15**.

Compound	Glide g-score	Prime-MM/GBSA ΔG_{bind} (in kcal/mol)	H-bond or Salt-bridge or CH···π/NH···π interactions	Residues within 4 Å
1	−7.91	−65	LysA30, LysA34, SerA47, SerA48	LysA32, AspA33, ValA35, PheA46, PheA49, SerA70, AspA71, TyrA75
2	−5.22	−61	AspA33, SerA47, SerA48, PheA49	LysA30, LysA34, ValA35, ArgA44, AlaA45, PheA46, LysA50
3	−6.61	−59	AspA33, SerA47, SerA48, LysA50	LysA30, LysA34, ValA35, ArgA44, AlaA45, PheA46, PheA49, ArgA52, SerA70, AspA71
4	−5.97	−58	LysA30, SerA47, SerA48, ArgA52, AspA71	AspA33, LysA34, ValA35, PheA46, PheA49, ThrA68, SerA70
5	−4.72	−56	AlaA45, PheA46, LysA50, AsnA51	LysA32, AspA33, LysA34, PheA43, ArgA44, SerA47, ValA53
6	−5.31	−59	AspA33, LysA34, SerA47	LysA30, LysA32, ValA35, GluA37, ArgA44, AlaA45, PheA46, SerA48, PheA49, LysA50
7	−5.20	−53	LysA30, LysA32, AspA33, LysA34, SerA47	GlnA31, ValA35, PheA46, SerA48, PheA49, TyrA75
8	−6.58	−54	LysA30, SerA47, SerA48, AspA71	AspA33, LysA34, ValA35, ArgA44, PheA46, PheA49, ArgA52, SerA70
9	−5.40	−54	AspA33, SerA47, SerA48	LysA30, ValA35, PheA46, PheA49, ArgA52, ThrA68, SerA70, AspA71
10	−5.38	−51	LysA34, PheA46, SerA47	LysA30, AspA33, ValA35, GluA37, ArgA44, SerA48, PheA49
11	−4.98	−57	LysA34, AlaA45, SerA47, LysA50	AspA33, PheA43, ArgA44, PheA46, AsnA51, ValA53
12	−4.52	−51	AspA33, SerA47, LysA50	LysA29, LysA32, LysA34, PheA43, ArgA44, AlaA45, PheA46, AsnA51, ValA53
13	−4.90	−50	LysA34, AlaA45, LysA50	AspA33, PheA43, ArgA44, PheA46, SerA47, AsnA51, ValA53
14	−4.78	−54	AlaA45, LysA50	LysA34, GluA37, GluA39, AsnA40, PheA43, ArgA44, PheA46, SerA47, AsnA51, ValA53
15	−3.97	−54	LysA34, LysA50	LysA32, AspA33, PheA43, ArgA44, AlaA45, PheA46, SerA47, AsnA51, ValA53

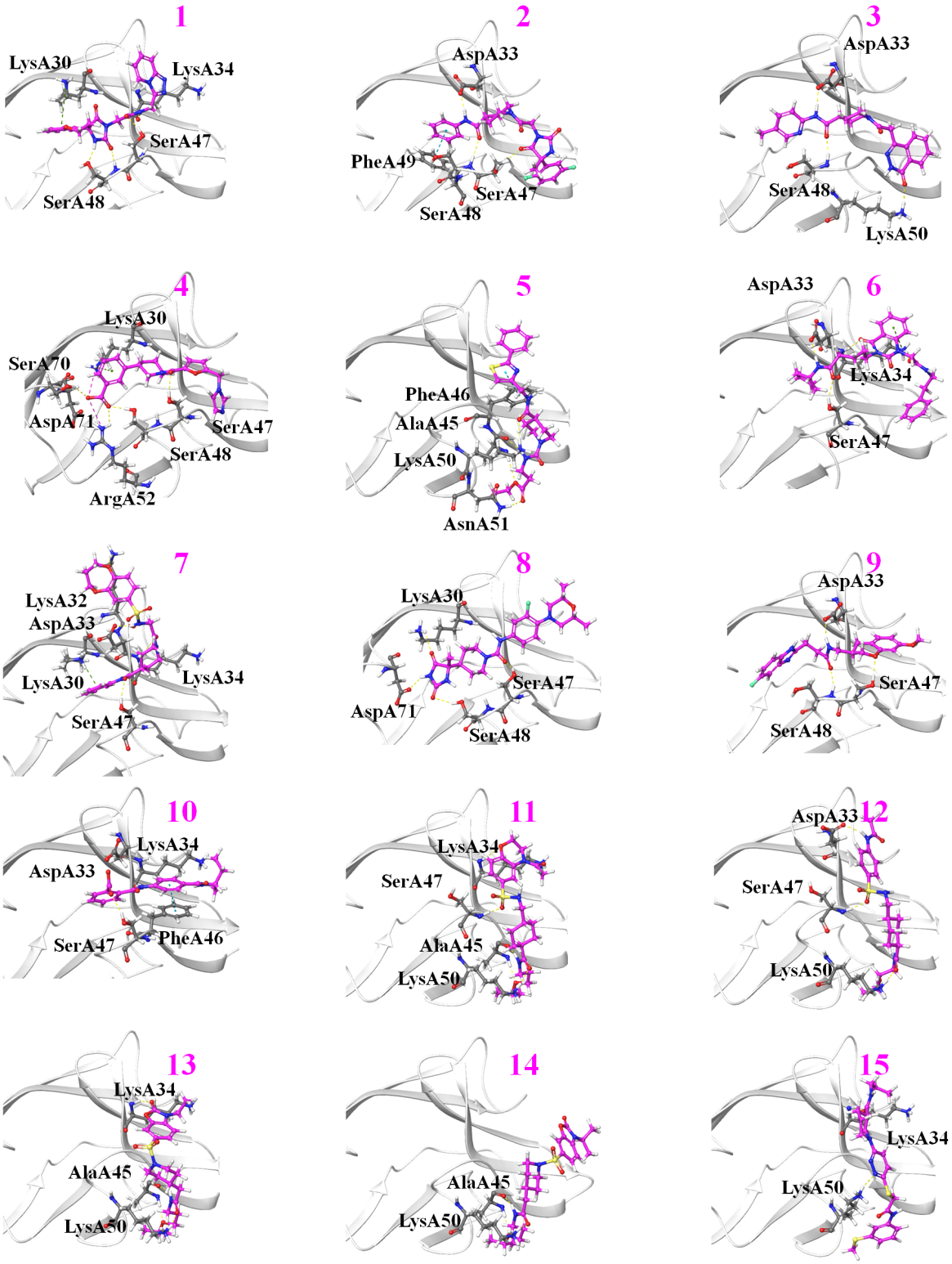

Figure 5. 3D interaction diagrams for the **compounds 1-15** after the Induced Fit Docking (IFD) with CD58. Hydrogen bonds are noted as yellow dotted lines, π···π interactions as blue dotted lined, NH···π/ CH···π interactions as green dotted lines, and salt bridges as magenta dotted lines.

MD Simulations of the Promising Hits

To simulate the binding of the selected compounds under physiological conditions, we performed MD simulations of the fifteen compounds selected after the IFD procedure, with a Prime-MM/GBSA ΔG_{bind} lower than −50 kcal/mol. The structural overlap of the top three conformations (after molecular docking) for each molecule is presented in Figure S23 and their respective binding affinities, calculated from Prime-MM/GBSA are indicated in Table S8. To identify the relevance of the binding conformation, the top three poses (in terms of binding affinity) (Figure S23) were considered for the MD simulations (50 ns each).

The system stability analysis was performed by the evaluation of the protein backbone and ligand RMSD over the simulation periods. As observed in Figure S24, the CD58 backbone exhibited an overall RMSD fluctuation of < 2 Å for all the ligand-bound systems during the three simulations (except for **13** and **14**). To select the promising candidate for the CD2-CD58 interaction inhibition, three criteria were taken into consideration: (i) a stabilized and low ligand RMSD for the bound conformation, (ii) a stabilized binding closer to the CD2-binding site (evaluated based on distance from AspA33 of CD58), and (iii) a comparative analysis with the initial binding conformations. For **compounds 1**, **3**, **4**, **8**, **9**, **10**, **12** (for Pose2 and Pose3), and **13** (for Pose2), the ligand RMSD values were lower than 3 Å from the initial conformation (Figure S25) with a fluctuation in the range of 1.0–1.5 Å, thereby indicating a stabilization of the ligand structure during the MD simulations. To evaluate the binding stability of the ligands at the CD2-binding site of CD58, the distance between the centroid of the bound ligand and AspA33 was analyzed (Figure S26). This distance was also compared to the initial distance between ligands and AspA33. The compounds which remained bound to CD58 at the CD2-binding site (distance from AspA33 < 8 Å) during all the three simulations are **1**, **4**, **8**, **9**, and **10**. Additionally, Pose2 of **compound 3**, Pose 3 of **compound 12**, and Pose 2 of **compound 13**, remained stably bound at the CD2-binding site of CD58 during the MD run. Although Pose3 of **compound 5** shifted to the position overlapping with the Pose1 (with a distance from AspA33 < 10 Å), **compound 5** was not selected as Pose1 was released from the binding site. The failure of Pose1 to stay in the same position reduced the confidence in this candidate. For the other simulations, the ligand was released from the binding site (indicated by an increased distance from AspA33). Based on these analyses, eight compounds, *i.e.,* **1**, **3**, **4**, **8**, **9**, **10**, **12**, and **13** (Figure 6) were proposed to exhibit stable binding to CD58 at the CD2-binding interface.

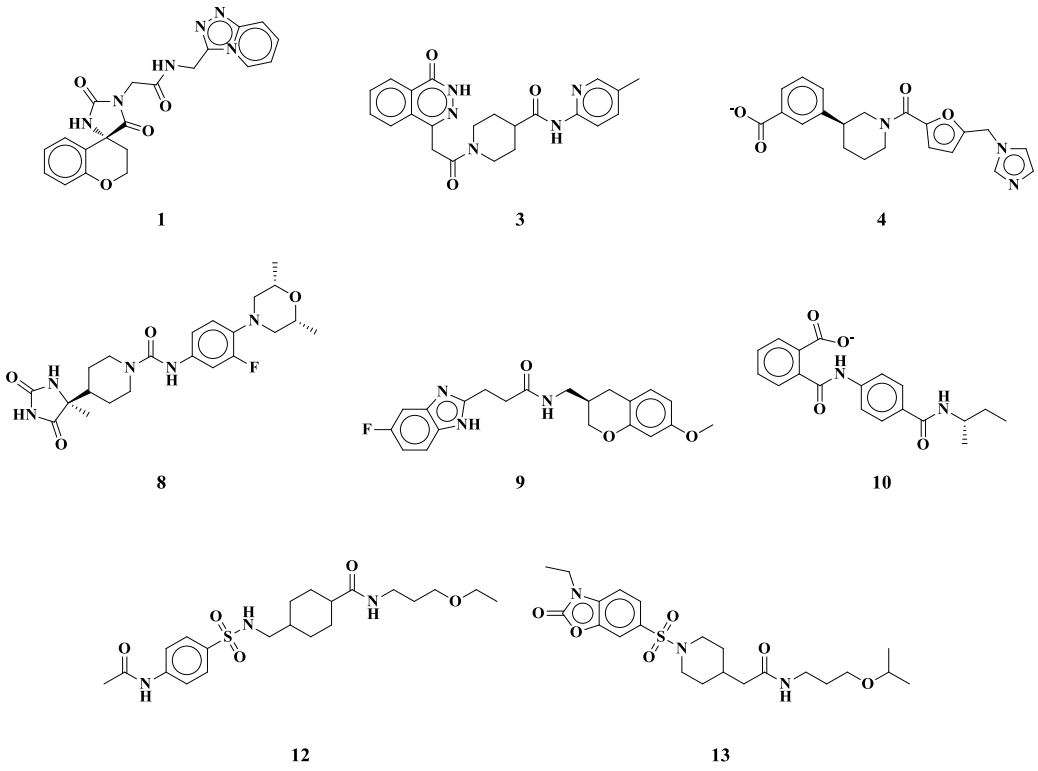

Figure 6. 2D structures of the eight selected molecules that exhibited stable binding to CD58 after the MD simulations.

The structural superimposition of the final positions of the ligands (after simulations of their three different starting points) for the eight compounds showed that at least two ligand conformations are highly overlapping for **compounds 1**, **4**, **8**, **9**, and **10** (Figure S27). For **compound 1**, the ligand stayed in a rather similar position for all three simulation systems (Figure S27A). Surprisingly, despite starting from highly overlapping conformations for **compounds 3**, **4**, **8**, **9**, and **10** (Figure S23), at least one of the three simulations behaved differently from the other two systems (Figure S27B-S27F). For **compounds 12** and **13**, the initial binding conformations were not overlapping (Figure S27G-S27H) and out of the three initial conformations, only one remained stably bound at the CD2-binding interface (close to AspA33) on CD58 for the two compounds. An additional constraint for the selection of the binding conformation was therefore imposed by analyzing their position on the CD2-CD58 interface (Figure 7). As noticed in the previous sections (Figures 3 and 4), the topographical N-terminal ends of CD2 and CD58 have a major contribution to the CD2-CD58 complex formation (Figure 7A). The MD simulation replicates which showed reproducible results (indicated by the overlapping position of the ligands after the dynamics in Figure S27A-S27D) were located at the topographical N-terminal end of CD58 for all selected compounds (Figure 7B-E), (except for **compound 10**), thus supporting their ability to block the

interactions between CD2 and CD58. For **compound 10**, Pose2 was located at the CD2-binding site (Figure 7G), although Pose1 and Pose3 behaved similarly (Figure S27F). Considering the pharmacological significance of this binding position, the final position from the MD simulations for Pose2 of **compound 10** was taken as the suitable binding conformation. After our careful analysis, the promising binding conformations proposed for the various **compounds** were Pose1 for **compounds 1**, **4**, and **9**, Pose 2 for **compounds 3**, **8**, **10**, and **13**, Pose 3 for **compound 12**.

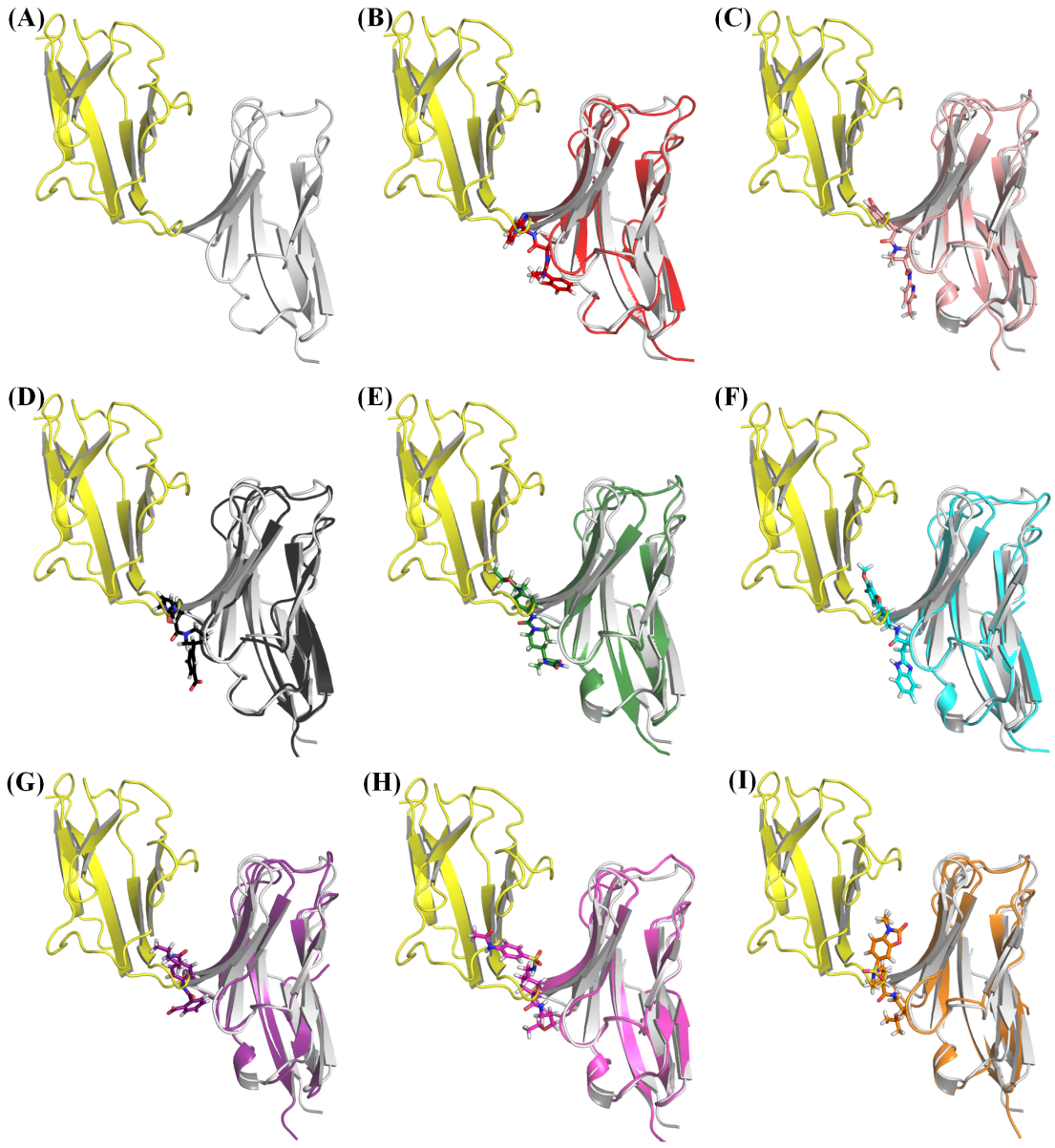

Figure 7. Structural superimposition of the CD2-CD58 complex (CD2 in yellow and CD58 in grey) (A)) with the optimized complexes (after MD) of the selected ligands and CD58 for **compound** (B) **1**, (C) **3**, (D) **4**, (E) **8**, (F) **9**, (G) **10**, (H) **12**, and (I) **13**.

The selected conformation and the ligand position obtained after MD simulations (Figure S28, shown in blue) were mostly in the same location as the one obtained after

molecular docking (Figure S28, in yellow). The residue-wise atomic fluctuation exhibited a similar pattern for the selected simulation system (corresponding to the most stable binding conformation) of the various complexes (Figure S29). The residues which exhibited the lowest fluctuation during the MD simulations include LysA30, AspA33, LysA34, ValA35, ArgA44, PheA46, SerA47, SerA48, PheA49, SerA69, SerA70, AspA71, and GluA72 (Figure S30). The lower fluctuation in these residues points towards their possible role in the binding of the selected compounds, as the complex formation results in a reduced degree of freedom for the residues located at the binding site. After the qualitative evaluation of the generated complexes and identification of the residues taking part in the complex formation, a quantitative evaluation of the complex formation was achieved through MM/GBSA ΔG_{bind} calculations for the eight compounds in the various ligand configurations.

For **compounds 1**, **3**, **4**, **8**, **9**, and **10**, at least two simulations exhibited a highly overlapping quantitative pattern of MM/GBSA ΔG_{bind} values over the period of 50 ns simulation (Figure S31). For **compounds 12** and **13**, we noted a significant difference in the binding energy for the three poses. The selected conformations (*i.e.*, Pose1 for **compounds 1**, **4**, and **9**, Pose 2 for **compounds 3**, **8**, **10**, and **13**, Pose 3 for **compound 12**) exhibited the highest binding affinity amongst the three simulations for each compound. Contrary to the CD2-CD58 complex formation, the binding energy of the selected compounds with CD58 was stabilized during the short period of simulation, which is attributed to the relatively overall smaller size of the bound ligands (as compared to CD2). Additionally, the selected compounds exhibited a higher degree of structural flexibility as compared to the cyclic peptides reported in literature,[21–25] which allowed the simulation algorithm to accommodate selected compounds on the CD2-binding interface. The MM/GBSA ΔG_{bind}, values averaged over the last 10 ns of MD simulations, of the selected binding conformations for various compounds are reported in Table 2. The order of binding strength for the selected molecules is **compounds 12** > **1** > **13** > **8** > **9** > **10** > **3** > **4**. The selected eight compounds exhibit heterocyclic rings (except **compounds 10** and **12**) and a bridging amide and/or sulphonamide functional group connecting two ring systems. The presence of multiple amide groups (or sulfonamide group) contributed positively to the binding energy of the selected compounds. Regarding **compounds 4** and **10**, the carboxylate functional group is exposed to solvent, and the contribution from ΔG_{solv} is, therefore, higher (numerically lower value) as compared to the other five non-ionic compounds. It was interesting to note that the presence of an amide

functional group (absent in **compound 4**) increased the contribution from the electrostatic energy (EEL).

The per-residue decomposition energy analysis for the various compounds helped to recognize the residues having a significant contribution to the complex formation (< −1 kcal/mol) (Figure S32). The important residues include LysA30, AspA33, LysA34, ValA35, PheA46, SerA47, SerA48, and PheA49. From those residues, LysA34, ValA35, PheA46, SerA47, and PheA49 were involved in ligand binding for at least three systems (indicated by multi-colored stacked bars in Figure 8A). The average number of hydrogen bonds between CD58 and the identified hits was calculated to be 2, 1, 1, 2, 2, 3, 2, and 2 for **compounds 1**, **3**, **4**, **8**, **9**, **10**, **12**, and **13**, respectively (Figure S33). The residues which were involved in the formation of the complex *via* H-bond formation are LysA30, AspA33, LysA34, SerA47, SerA48, and PheA49 (Figure 8B). The side chains of the residues participating in interactions with the selected ligands adopted a different configuration to accommodate the ligands more comfortably at the CD2-interacting surface (Figure S34). The p*Ka* values were predicted for CD58 by using the centroid of each cluster (based on interfacial residues, Figure S9) using the PDB2PQR server (http://server.poissonboltzmann.org/pdb2pqr).[56] The results indicated that despite the change in the side change orientations in CD58, there were no significant differences between the predicted p*Ka* values for the various amino acids (Figure S35). To evaluate the effect of the force field on the binding of ligands to CD58, we performed a comparative analysis for the compound **1** by applying the GAFF2,[31] and FF14SB force field.[32] Although the calculated affinity with recent force fields was weaker than that with GAFF[29] and FF99SB,[30] the pattern of MM/GBSA ΔG_{bind} values was comparable with the two force fields (Figure S36). A comparative analysis of the important residues interacting with the previously reported cyclic peptides[24–26] and the identified hits from the current study was performed thereafter. For the cyclic peptides, GluA25, LysA29, AspA33, LysA34, and ArgA44 from CD58 are reported to participate in the complex formation.[26] LysA34 was identified as a strong anchor point in the CD58-cyclic peptide complex.[26] Interestingly, AspA33 and LysA34 also played a key role in the binding of the selected hits in this work, which increases the confidence in the inhibitory potential of the selected hits. As already mentioned, the reported cyclic peptide inhibitors[24,25] were observed to move from the initial and expected site of interaction in CD58,[26] which can be the reason for a lower overlap in the interacting residue for the identified ligands reported herein.

Table 2. MM/GBSA ΔG_{bind} values in kcal/mol calculated after the 50 ns MD simulations (averaged over the last 10 ns) for the selected compounds.*

Compound	VDW	EEL	EGB	ESURF	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
1	-36	-30	43	-5	-66	37	-28 ± 3
3	-29	-19	34	-4	-48	31	-17 ± 3
4	-27	47	-33	-4	20	-37	-17 ± 3
8	-26	-26	33	-4	-52	29	-23 ± 4
9	-26	-25	36	-3	-51	33	-18 ± 3
10	-19	-7	11	-3	-26	8	-18 ± 4
12	-27	-30	32	-4	-57	28	-29 ± 4
13	-28	-16	24	-4	-44	20	-24 ± 4

*The meaning of the different terms used in this table is as follows: **VDW** = van der Waals energy as calculated by the MM force field. **EEL** = electrostatic energy is calculated by the MM force field. **EGB** = the electrostatic contribution to the solvation free energy calculated by GB. **ESURF** = nonpolar contribution to the solvation free energy calculated by an empirical model. **ΔG_{gas}** = total gas phase energy, *i.e.*, the sum of the van der Waals and electrostatic energy from MM. **ΔG_{solv}** = total free energy of solvation, *i.e.*, the sum of the electrostatic and nonpolar contributions from solvation. **ΔG_{bind}** = final estimated binding free energy calculated from the terms above (kcal/mol).

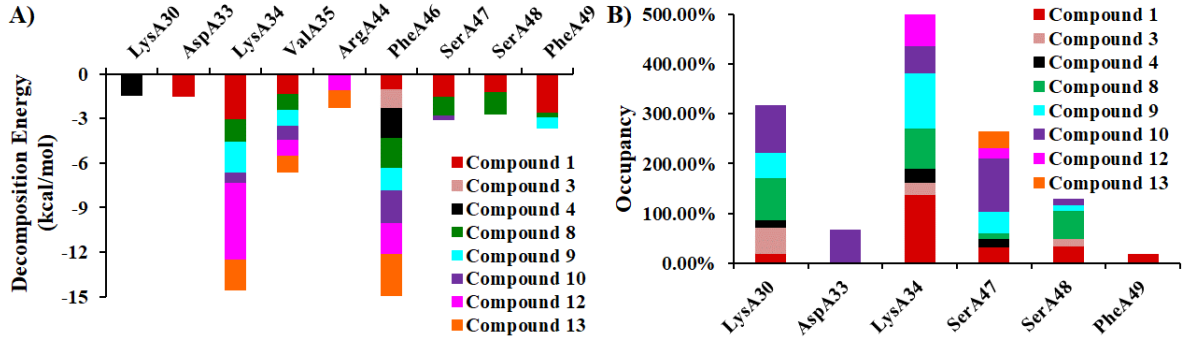

Figure 8. Identification of the interacting residues for the eight selected compounds after the MD simulations with the help of the (A) decomposition energy analysis and (B) the H-bond occupancy analysis, for which the cumulative percentage of frames exhibiting the presence of H-bond with each residue for all the compounds is provided.

From the detailed analyses of MD simulations, it can thus be suggested that **compounds 1**, **3**, **4**, **8**, **9**, **10, 12**, and **13** exhibited the binding characteristics (including binding position, binding energy, and the residues involved in the interactions) which make them promising hits as inhibitors of CD2 interactions with CD58. The various molecular properties of the selected compounds, which are known to affect the pharmacokinetics of small molecules, were within the acceptable range, *i.e.*, molecular weight < 500 Da, LogP < 5, HBA < 10, HBD < 5, TPSA < 140 Å^2 (Table S9). Source libraries and commercial details of these compounds are provided in Table S10. The selected compounds offer the opportunity of developing non-peptidic inhibitors of CD2-CD58 interactions and can be evaluated experimentally to ascertain the findings from our work.

Conclusions

The interaction between CD2 and CD58 is the preliminary phenomenon in the physiological immune response mediated *via* T-cells and APCs. Inhibition of CD2-CD58 interactions is established as a suitable strategy to develop immunosuppressive agents, which have a wide range of therapeutic applications. With this information in hand, we performed a systematic virtual screening for the identification of CD58-binding agents which have the potential to inhibit the interaction of CD2 with CD58. In the first step, MD simulations (500 ns, in triplicate) were performed for the CD2-CD58 complex. After ensuring the stabilization of the systems and reproducibility of the results, the residues which stably contributed to the CD2 binding by more than −1 kcal/mol were identified. These include LysA32, AspA33, ArgA44, and PheA46. A pharmacophore map was created with these residues by incorporating nine structural features, *i.e.*, two HBA (A_1 and A_4), three HBD (D_6, D_7, and D_9), two hydrophobic groups (H_{10} and H_{11}), one negative functional group (N_{12}), and one aromatic group (R_{15}). Screening against the developed pharmacophore resulted in 5,882 compounds out of the 227,228 molecules, which exhibited at least six out of nine features and were taken for the next step property-based filtering. Druglikeness and ADME evaluation resulted in funneling the number of selected compounds to 343 molecules.

Molecular docking based screening was next performed using induced fit docking, and subsequent Prime MM/GBSA ΔG_{bind} calculations. From the screened set, fifteen compounds exhibited a binding affinity equivalent to −50 kcal/mol, which were taken for the next step, *i.e.*, MD simulations (50 ns, in triplicate). A systematic analysis of the dynamics helped to identify eight compounds exhibiting reproducible and stable binding with CD58. The selected **compounds 1**, **3**, **4**, **8**, **9**, **10, 12**, and **13**, exhibited MM/GBSA ΔG_{bind} of −28, −17, −17, −23, −18, −18, −29, and −24 kcal/mol, respectively. The residues which are involved in the interactions with identified hits include LysA30, AspA33, LysA34, ValA35, PheA46, SerA47, SerA48, and PheA49 are located at the CD2-binding interface of CD58. The interaction of the selected hits with the topographical N-terminal end of CD58 at the CD2-binding interface further supports their candidature as CD58-binding agents, and hence an experimental evaluation can be followed. Let us stress that the selected compounds have passed through the various applied ADMET filters. The hits identified in our work open the opportunity for the design and identification of non-peptidic inhibitors of CD2-CD58 interactions.

Acknowledgments

NT thanks the Région Pays de la Loire and the Centre National de la Recherche Scientifique (CNRS) for the financial support during her post-doctoral research within the PIRAMID and MimBreg projects. ADL acknowledges the Région Pays de la Loire for financial support within the framework of PIRAMID and MiM-Breg project. This research used computational resources of (i) CCIPL (Centre de Calcul Intensif des Pays de Loire). Funding was provided by the Wallonie-Bruxelles International WBI (PHC Tournesol DoIFAD) and the Belgian National Foundation for Scientific Research (FNRS), by the French Ministry of Foreign and European Affairs, and by the Ministry of Higher Education and Research, in the framework of the Hubert Curien partnerships (PHC Tournesol #40638PL).

Conflicts of interest

The authors declare no conflict of interest.

List of Abbreviations

ΔG_{bind}	: Binding energy
Å	: Angstrom
APCs	: Antigen-presenting cells
CD2	: Cluster of differentiation 2
DruLiTo	: Drug Likeness Tool
H-bonds	: Hydrogen bonds
IFD	: Induced fit docking
MD	: Molecular dynamics
MM/GBSA	: Molecular Mechanics-Generalized Born Surface Area
PME	: Particle Mesh Ewald
PDB	: RCSB/Protein Data Bank
QED	: Quantitative estimation of druglikeness
RMSD	: Root mean square deviation
SP	: Standard Precision
uwQED	: Unweighted quantitative estimation of druglikeness
VMD	: Visual Molecular Dynamics
wQED	: Weighted quantitative estimation of druglikeness
XP	: eXtra Precision

References

1 R. Medzhitov and C. A. Janeway Jr., Innate immunity: Impact on the adaptive immune response, *Curr. Opin. Immunol.*, 1997, **9**, 4–9.

2 J. Huang, C. Meyer and C. Zhu, T cell antigen recognition at the cell membrane, *Mol. Immunol.*, 2012, **52**, 155–164.

3 T. A. Springer, Adhesion receptors of the immune system, *Nature*, 1990, **346**, 425–434.

4 M. C. Montoya, D. Sancho, M. Vicente-Manzanares and F. Sánchez-Madrid, Cell adhesion and polarity during immune interactions, *Immunol. Rev.*, 2002, **186**, 68–82.

5 P. Wang, B. T. Qi, P. Chen, L. J. He, J. Li, Y. Q. Ji and M. Xie, CD58 expression of liver tissue in patients with chronic hepatitis B virus infection, *Chin. Med. J. (Engl).*, 2008, **121**, 557–560.

6 L. Sheng, J. Li, B. T. Qi, Y. Q. Ji, Z. J. Meng and M. Xie, Investigation on correlation between expression of CD58 molecule and severity of hepatitis B, *World J. Gastroenterol.*, 2006, **12**, 4237–4240.

7 T. Shao, W. Shi, J. Y. Zheng, X. X. Xu, A. F. Lin, L. X. Xiang and J. Z. Shao, Costimulatory function of Cd58/Cd2 interaction in adaptive humoral immunity in a zebrafish model, *Front. Immunol.*, 2018, **9**, 1204.

8 C. F. Mojcik and E. M. Shevach, Adhesion molecules: A rheumatologic perspective, *Arthritis Rheum.*, 1997, **40**, 991–1004.

9 S. Raychaudhuri, B. P. Thomson, E. F. Remmers, S. Eyre, A. Hinks, C. Guiducci, J. J. Catanese, G. Xie, E. A. Stahl, R. Chen, L. Alfredsson, C. I. Amos, K. G. Ardlie, A. Barton, J. Bowes, N. P. Burtt, M. Chang, J. Coblyn, K. H. Costenbader, L. A. Criswell, J. B. A. Crusius, J. Cui, P. L. de Jager, B. Ding, P. Emery, E. Flynn, P. Harrison, L. J. Hocking, T. W. J. Huizinga, D. L. Kastner, X. Ke, F. A. S. Kurreeman, A. T. Lee, X. Liu, Y. Li, P. Martin, A. W. Morgan, L. Padyukov, D. M. Reid, M. Seielstad, M. F. Seldin, N. A. Shadick, S. Steer, P. P. Tak, W. Thomson, A. H. M. van der Helm-Van Mil, I. E. van der Horst-Bruinsma, M. E. Weinblatt, A. G. Wilson, G. J. Wolbink, P. Wordsworth, D. Altshuler, E. W. Karlson, R. E. M. Toes, N. de Vries, A. B. Begovich, K. A. Siminovitch, J. Worthington, L. Klareskog, P. K. Gregersen, M. J. Daly and R. M. Plenge, Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk, *Nat. Genet.*, 2009, **41**, 1313–1320.

10 J. Lerut, V. Van Thuyne, J. Mathijs, J. Lemaire, S. Talpe, F. Roggen, O. Ciccarelli, M. Zuckermann, P. Goffette, J. Hope, P. Gianello, H. Bazin, A. Cornet, J. Rahier and D. Latinne, Anti-CD2 monoclonal antibody and tacrolimus in adult liver transplantation, *Transplantation*, 2005, **80**, 1186–1193.

11 E. Lee, S. J. Lee, J. E. Koskimaki, Z. Han, N. B. Pandey and A. S. Popel, Inhibition of breast cancer growth and metastasis by a biomimetic peptide, *Sci. Rep.*, 2014, **4**, 7139.

12 J. A. Gollob, J. Li, H. Kawasaki, J. F. Daley, C. Groves, E. L. Reinherz and J. Ritz, Molecular interaction between CD58 and CD2 counter-receptors mediates the ability of monocytes to augment T cell activation by IL-12, *J. Immunol.*, 1996, **157**, 1886–1893.

13 A. Aruffo and D. Hollenbaugh, Therapeutic intervention with inhibitors of co-

stimulatory pathways in autoimmune disease, *Curr. Opin. Immunol.*, 2001, **13**, 683–686.

14 D. Przepiorka, G. L. Phillips, V. Ratanatharathorn, M. Cottler-Fox, L. H. Sehn, J. H. Antin, D. LeBherz, M. Awwad, J. Hope and J. B. McClain, A phase II study of BTI-322, a monoclonal anti-CD2 antibody, for treatment of steroid-resistant acute graft-versus-host disease, *Blood*, 1998, **92**, 4066–4071.

15 L. Branco, P. Barren, S. Y. Mao, D. Pfarr, R. Kaplan, C. Postema, S. Langermann, S. Koenig and S. Johnson, Selective deletion of antigen-specific, activated T cells by a humanized MAB to CD2 (MEDI-507) is mediated by NK cells, *Transplantation*, 1999, **68**, 1588–1596.

16 J. Liu, V. T. K. Chow and S. D. S. Jois, A novel, rapid and sensitive heterotypic cell adhesion assay for CD2-CD58 interaction, and its application for testing inhibitory peptides, *J. Immunol. Methods*, 2004, **291**, 39–49.

17 J. H. Wang, A. Smolyar, K. Tan, J. H. Liu, M. Kim, Z. Y. J. Sun, G. Wagner and E. L. Reinherz, Structure of a heterophilic adhesion complex between the human CD2 and CD58 (LFA-3) counterreceptors, *Cell*, 1999, **97**, 791–803.

18 M. V. Bayas, K. Schulten and D. Leckband, Forced detachment of the CD2-CD58 complex, *Biophys. J.*, 2003, **84**, 2223–2233.

19 X. Wang, C. G. Ji and J. Z. H. Zhang, Glycosylation modulates human CD2-CD58 adhesion via conformational adjustment, *J. Phys. Chem. B*, 2015, **119**, 6493−6501.

20 Z. Y. J. Sun, V. Dötsch, M. Kim, J. Li, E. L. Reinherz and G. Wagner, Functional glycan-free adhesion domain of human cell surface receptor CD58: Design, production and NMR studies, *EMBO J.*, 1999, **18**, 2941–2949.

21 J. Liu, J. Ying, V. T. K. Chow, V. J. Hruby and S. D. Satyanarayanajois, Structure-activity studies of peptides from the 'hot-spot' region of human CD2 protein: Development of peptides for immunomodulation, *J. Med. Chem.*, 2005, **48**, 6236–6249.

22 J. Liu, C. Li, S. Ke and S. D. Satyanarayanajois, Structure-based rational design of β-hairpin peptides from discontinuous epitopes of Cluster of Differentiation 2 (CD2) protein to modulate cell adhesion interaction, *J. Med. Chem.*, 2007, **50**, 4038–4047.

23 S. Giddu, V. Subramanian, H. S. Yoon and S. D. Satyanarayanajois, Design of β-hairpin peptides for modulation of cell adhesion by β-turn constraint, *J. Med. Chem.*, 2009, **52**, 726–736.

24 A. Gokhale, T. K. Weldeghiorghis, V. Taneja and S. D. Satyanarayanajois, Conformationally constrained peptides from CD2 to modulate protein-protein interactions between CD2 and CD58, *J. Med. Chem.*, 2011, **54**, 5307–5319.

25 A. S. Gokhale, R. Sable, J. D. Walker, L. McLaughlin, K. G. Kousoulas and S. D. Jois, Inhibition of cell adhesion and immune responses in the mouse model of collagen-induced arthritis with a peptidomimetic that blocks CD2-CD58 interface interactions, *Biopolymers*, 2015, **104**, 733–742.

26 L. Leherte, A. Petit, D. Jacquemin, D. P. Vercauteren and A. D. Laurent, Investigating cyclic peptides inhibiting CD2–CD58 interactions through molecular dynamics and molecular docking methods, *J. Comput. Aided. Mol. Des.*, 2018, **32**, 1295–1313.

27 D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, V. W.

D. Cruzeiro, T. A. Darden, R. E. Duke, D. Ghoreishi, M. K. Gilson, H. Gohlke, A. W. Goetz, D. Greene, R. Harris, N. Homeyer, S. Izadi, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. J. Mermelstein, K. M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R. C. Walker, J. Wang, H. Wei, R. M. Wolf, X. Wu, L. Xiao, D. M. York and P. A. Kollman, *Univ. California, San Fr.*, 2018.

28 D. C. Bas, D. M. Rogers and J. H. Jensen, Very fast prediction and rationalization of pKa values for protein-ligand complexes, *Proteins Struct. Funct. Genet.*, 2008, **73**, 765–783.

29 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, Development and testing of a general Amber force field, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

30 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, Comparison of multiple amber force fields and development of improved protein backbone parameters, *Proteins Struct. Funct. Genet.*, 2006, **65**, 712–725.

31 D. Vassetti, M. Pagliai and P. Procacci, Assessment of GAFF2 and OPLS-AA general force fields in combination with the water models TIP3P, SPCE, and OPC3 for the solvation free energy of druglike organic molecules, *J. Chem. Theory Comput.*, 2019, **15**, 1983–1995.

32 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, ff14SB: Improving the accuracy of protein side chain and backbone Parameters from ff99SB, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.

33 D. R. Roe and T. E. Cheatham III, PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.

34 W. Humphrey, A. Dalke and K. Schulten, VMD: Visual molecular dynamics, *J. Mol. Graph.*, 1996, **14**, 33–38.

35 S. Genheden and U. Ryde, The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities, *Expert Opin. Drug Discov.*, 2015, **10**, 449–461.

36 A. R. N. Arulanandam, J. M. Withka, D. F. Wyss, G. Wagner, A. Kister, P. Pallai, M. A. Recny and E. L. Reinherz, The CD58 (LFA-3) binding site is a localized and highly charged surface area on the AGFCC'C'' face of the human CD2 adhesion domain, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 11613–11617.

37 S. L. Dixon, A. M. Smondyrev, E. H. Knoll, S. N. Rao, D. E. Shaw and R. A. Friesner, PHASE: A new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results, *J. Comput. Aided. Mol. Des.*, 2006, **20**, 647–671.

38 S. L. Dixon, A. M. Smondyrev and S. N. Rao, PHASE: A novel approach to pharmacophore modeling and 3D database searching, *Chem. Biol. Drug Des.*, 2006, **67**, 370–372.

39 Phase, Schrödinger LLC NY, 2019.

40 G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju and W. Sherman, Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments, *J. Comput. Aided. Mol. Des.*, 2013, **27**, 221–234.

41 LigPrep, Schrödinger LLC NY, 2019.

42 J. C. Shelley, A. Cholleti, L. L. Frye, J. R. Greenwood, M. R. Timlin and M. Uchimaya, Epik: A software program for pKa prediction and protonation state generation for drug-like molecules, *J. Comput. Aided. Mol. Des.*, 2007, **21**, 681–691.

43 J. R. Greenwood, D. Calkins, A. P. Sullivan and J. C. Shelley, Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution, *J. Comput. Aided. Mol. Des.*, 2010, **24**, 591–604.

44 *Schrödinger LLC NY*, 2019.

45 K. S. Watts, P. Dalal, R. B. Murphy, W. Sherman, R. A. Friesner and J. C. Shelley, ConfGen: A conformational search method for efficient generation of bioactive conformers, *J. Chem. Inf. Model.*, 2010, **50**, 534–546.

46 ConfGen, Schrödinger, LLC, New York, NY, 2019.

47 DruLiTo (http://www.niper.gov.in/pi_dev_tools/DruLiToWeb/DruLiTo_index.html).

48 C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Deliv. Rev.*, 1997, **23**, 3–25.

49 A. K. Ghose, V. N. Viswanadhan and J. J. Wendoloski, A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases, *J. Comb. Chem.*, 1999, **1**, 55–68.

50 D. F. Veber, S. R. Johnson, H. Y. Cheng, B. R. Smith, K. W. Ward and K. D. Kopple, Molecular properties that influence the oral bioavailability of drug candidates, *J. Med. Chem.*, 2002, **45**, 2615–2623.

51 T. I. Oprea, Property distribution of drug-related chemical databases, *J. Comput. Aided. Mol. Des.*, 2000, **14**, 251–264.

52 U. Norinder and M. Haeberlein, Computational approaches to the prediction of the blood-brain distribution, *Adv. Drug Deliv. Rev.*, 2002, **54**, 291–313.

53 G. R. Bickerton, G. V. Paolini, J. Besnard, S. Muresan and A. L. Hopkins, Quantifying the chemical beauty of drugs, *Nat. Chem.*, 2012, **4**, 90–98.

54 QikProp, Schrödinger, LLC, New York, NY, 2019.

55 Protein Preparation Wizard, Schrödinger, LLC, New York, NY, 2019.

56 T. J. Dolinsky, J. E. Nielsen, J. A. McCammon and N. A. Baker, PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations, *Nucleic Acids Res.*, 2004, **32**, W665–W667.

57 R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, *J. Med. Chem.*, 2004, **47**, 1739–1749.

58 T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard and J. L. Banks, Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening, *J. Med. Chem.*, 2004, **47**, 1750–1759.

59 R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin and D. T. Mainz, Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, *J. Med. Chem.*, 2006, **49**, 6177–6196.

60 Glide, version 5.7, Schrödinger, LLC, New York, NY, 2019.

61 W. Sherman, T. Day, M. P. Jacobson, R. A. Friesner and R. Farid, Novel procedure for modeling ligand/receptor induced fit effects, *J. Med. Chem.*, 2006, **49**, 534–553.

62 W. Sherman, H. S. Beard and R. Farid, Use of an induced fit receptor structure in virtual screening, *Chem. Biol. Drug Des.*, 2006, **67**, 83–84.

63 Induced Fit Docking Protocol, Schrödinger LLC NY, 2019.

64 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, *J. Am. Chem. Soc.*, 1996, **118**, 11225–11236.

Structure-Based Identification of Inhibitors Disrupting the CD2-CD58 Interactions

Neha Tripathi,[1] Laurence Leherte,[2,3] Daniel P. Vercauteren,[2] Adèle D. Laurent[1]**

[1]Université de Nantes, CNRS, CEISAM UMR 6230, F-44000 Nantes, France.

[2]Unit of Theoretical and Structural Physical Chemistry, Department of Chemistry, NAmur Research Institute for LIfe Sciences (NARILIS), NAmur MEdicine & Drug Innovation Center (NAMEDIC), Namur Institute of Structured Matter (NISM), University of Namur, Rue de Bruxelles 61, B-5000 Namur (Belgium).

[3]Laboratory of Structural Biological Chemistry, University of Namur, Rue de Bruxelles 61, B-5000 Namur (Belgium).

*Corresponding Authors: Neha Tripathi, Email: Neha.Tripathi@univ-nantes.fr and Adèle D. Laurent, Email : Adele.Laurent@univ-nantes.fr

Supporting Information

Computational Methodology

Molecular Dynamics simulations

The crystallographic structure of the complex CD2-CD58 is available in the RCSB/Protein Data Bank (PDB ID: 1QA9).[1] To evaluate the dynamics of the complex and to identify the residues which are critical for its formation, MD simulations were performed using the AMBER18 package.[2] Additionally, complexes between CD58 and hits obtained after virtual screening (discussed in subsequent sections) were also considered through MD simulations. The systems preparation was performed with the help of Amber tools.[2] Ligands and proteins were handled using the General Amber Force Field (GAFF)[3] and the Amber ff99SB force field,[4] respectively. When needed, the atomic charges of the small molecules were derived using the AM1-bcc method (semi-empirical with bond charge correction)[5] of the antechamber module from Amber tools.[2] Explicit water molecules (using TIP3P water model)[6] were considered for solvation of the systems (with a cubic box with an edge not closer than 15 Å to the protein) and counter ions were added to neutralize the overall charge on the generated assemblies. Additionally, a required number of Na^+ and Cl^- ions were included to maintain an ionic concentration of 0.15 M (isotonic to body fluids) (Table S1).

Table S1. Details of the generated assemblies for MD simulations.

Sl. No.	System	Number of Na^+	Number of Cl^-	Number of Solvent Molecules	Total Number of atoms
1	CD2-CD58 Complex	68	64	15,710	50,490
2	*Apo*-CD2	33	32	11,100	35,071
3	*Apo*-CD58	23	20	8,864	28,157

The generated assemblies were submitted to a three-step minimization, in which restraint weights of 3.0 and 1.0 kcal/mol/A^2 were applied at the first two steps, respectively, on the protein and ligands, whereas solvent and ions were minimized without any restraint. The third step was an unrestrained minimization of the entire system. Thereafter gradual heating (from 0 to 300 K, under NVT ensemble) was performed. This was followed by a three-step density equilibration (under NPT ensemble), in which the restraint force constant was decreased from 2.0 to 1.0 to 0.0 kcal/mol/A^2, on the protein and ligands. Finally, an equilibration (1 ns under NPT, at 310 K and 1 atm pressure) and production run (1000 ns under NPT) were performed (pressure relaxation time = 2.0 ps, time step = 2 fs, and cutoff distance for non-bonded interactions = 12 Å). The particle-Mesh Ewald (PME) method[7] was employed for the treatment of long-range electrostatic interactions and the SHAKE algorithm[8] was used to constraint all the covalent bonds involving hydrogen atoms. The bulk

effect was simulated by enabling periodic boundary conditions. All the simulations were run in triplicate to evaluate the reproducibility of the results. The result analysis was performed with the help of the Ptraj module[9] of the Amber tools set[2] and Visual Molecular Dynamics software (VMD).[10] The replicate simulations were submitted to a combined clustering analysis (using the Ptraj module)[9] by employing the hierarchical agglomerative (bottom-up) approach as the clustering algorithm (maximum number of clusters: 5) and considering the best-fit coordinate root mean square deviation (RMSD) between all the heavy atoms as clustering parameter. The calculation of the binding energy (ΔG_{bind}) values was performed using the MM/GBSA method.[11]

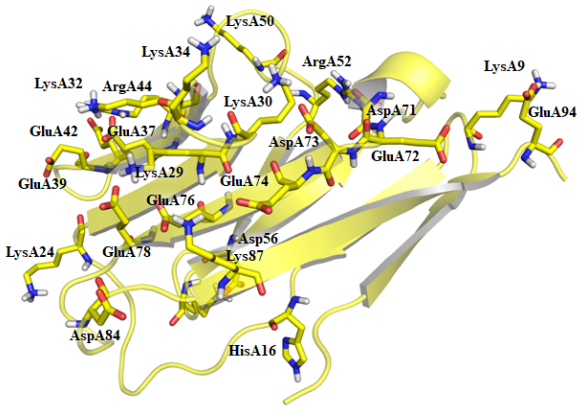

Figure S9. The ionizable residues in the CD58 structure and their states predicted from PROPKA, which were subsequently utilized for molecular docking and molecular dynamics.

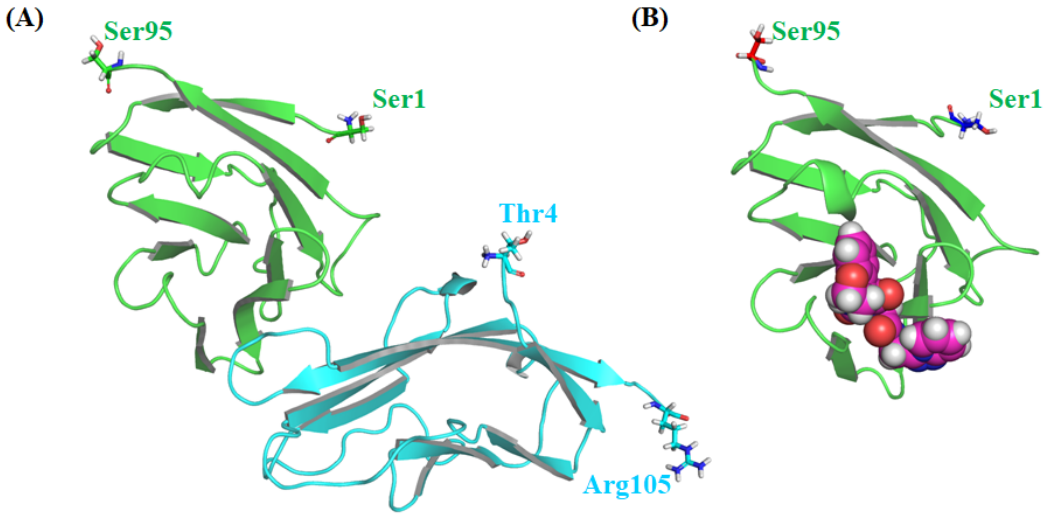

Figure S10. Location of the N-terminal and C-terminal residues (A) in the CD2-CD58 complex (CD2 in cyan and CD58 in green, PDB ID: 1QA9), and (B) with respect to the expected ligand-binding site.

Table S2. Molecules considered for virtual screening along with their source libraries.

Sl. No.	Library	Number of compounds
1	ASINEX	11,377
2	Chembridge	100,000
3	Chemdiv	12,995
4	Enamine	65,512
5	LifeChemicals	22,944
6	Maybridge	14,400
	Total	**227,228**

Table S3. QikProp properties considered for the property based virtual screening.

Molecular Property calculated using QikProp	
Molecular weight	Dipole moment (D)
Total solvent accessible surface area (SASA) ($Å^2$)	Hydrophobic SASA ($Å^2$)
Hydrophilic SASA ($Å^2$)	Carbon Pi SASA ($Å^2$)
Weakly polar SASA molecular volume ($Å^3$)	van der Waals polar surface area (PSA) ($Å^2$)
Number of rotatable bonds	Number of hydrogen bond donors
Number of hydrogen bond acceptors	Globularity (sphere = 1)
Ionization potential (eV)	Electron affinity (eV)
QP polarizability ($Å^3$)	QP LogP (in various solvents)
QP LogS for aqueous solubility	Conformation independent QP LogS
QP $LogK_{hsa}$ as a measure of serum protein binding	QP LogBB for brain/blood
Number of primary metabolites (#metabol)	Predicted CNS activity (-- to ++)
Log IC_{50} HERG K^+ channel blockage (QplogHERG)	Apparent Caco-2 permeability (nm/sec) (QpPcaco)
Apparent MDCK permeability (nm/sec)	QP LogKp for skin permeability
Max transdermal transport rate (J_m)	Lipinski rule of 5 violations
Jorgensen rule of 3 violations	Qualitative model for human oral absorption
% Human oral absorption in gastrointestinal tract (±20 %)	

Results and Discussion

Structural information

Table S4. Structural information available in RCSB/PDB for CD2 and CD58.

PDB ID	Macromolecular Structure	Number of Residues	Resolution	R-Value Free	R-value Work
1HNF	*apo*-CD2	182	2.5 Å	-NA-	0.193
1CCZ	*apo*-CD58	171	1.8 Å	0.243	0.202
1QA9	CD2-CD58 complex	102-95 (CD2-CD58)	3.2 Å	0.281	0.223

Table S5. List of atom pairs with close contacts in the crystal structure (PDB ID: 1QA9).[1] Chain A refers to CD58 (highlighted in green) and Chain B refers to CD2 in the dimer used for the molecular modeling studies. The intermolecular close contacts are highlighted in yellow and are shown in bold.

Sl. No.	Atom1	Atom2	Interatomic distance (Å)	Clash overlap (Å)
1	ThrA83@H	IleA82@HG12	1.02	1.13
2	IleA82@HB	SerA79@O	1.57	1.04
3	ThrA83@N	IleA82@HG12	1.74	1.01
4	**LysA34@NZ**	**ArgB48@HH21**	**1.70**	**0.88**
5	ValB93@HB	IleB85@HG22	1.56	0.87
6	TyrA65@HA	AsnA12@HB3	1.59	0.85
7	**LysA34@HZ3**	**ArgB48@HH21**	**1.24**	**0.83**
8	**LysA34@NZ**	**ArgB48@NH2**	**2.30**	**0.80**
9	ValB93@HB	IleB85@CG2	2.11	0.80
10	LysB96@NZ	ProB22@HG2	1.97	0.79
11	MetA86@HE1	ValA17@HG13	1.72	0.72
12	GlyB66@O	IleB21@HD11	1.89	0.71
13	GluB52@HB3	LysB51@O	1.89	0.70
14	GlnA31@HB2	LysA30@O	1.92	0.69
15	AlaB67@HB3	LeuB63@HB2	1.74	0.69
16	**LysA34@HZ3**	**ArgB48@NH2**	**1.89**	**0.66**
17	AspA84@H	IleA82@CD1	2.09	0.66
18	AspA84@HA	IleA82@HD11	1.77	0.66
19	AsnB65@HB2	LysB64@O	1.99	0.63
20	**LysA34@HZ1**	**ArgB48@NH2**	**1.97**	**0.62**
21	LysA87@HG2	GluA76@HB2	1.82	0.62
22	LysA87@O	SerA2@HA	2.00	0.62
23	LysB96@HZ2	ProB22@HG2	1.65	0.61
24	ThrA83@N	IleA82@CG1	2.51	0.60
25	ArgB48@HE	AspB32@HB2	1.66	0.60
26	ThrA14@H	ValA13@HG12	1.67	0.59
27	IleA82@HG22	ProA80@O	2.02	0.58
28	SerA79@HB2	ValA26@HG22	1.86	0.57

29	IleA82@H	ProA80@C	2.05	0.57
30	AspA84@N	IleA82@CD1	2.68	0.57
31	IleB102@O	IleB102@HG13	2.04	0.57
32	LysB69@N	LeuB68@HD22	2.18	0.57
33	IleB85@HD11	IleB33@HD13	1.85	0.57
34	AspA84@CA	IleA82@HD11	2.35	0.56
35	ValA92@HG11	SerA69@HA	1.88	0.56
36	LeuB100@HD23	AspB77@O	2.06	0.56
37	AspA84@N	IleA82@HD11	2.22	0.55
38	TyrA91@HA	GluA72@HG3	1.88	0.55
39	SerA41@HB2	AsnA40@O	2.08	0.54
40	ArgB48@HA	AspB31@O	2.08	0.54
41	LysB69@O	TyrB60@HA	2.07	0.54
42	ArgA52@H	LysA50@C	2.11	0.53
43	AspA84@H	IleA82@HD13	1.75	0.52
44	AspB99@HB3	GluB8@HA	1.92	0.52
45	GlyA93@HA3	GlyA7@O	2.11	0.51
46	AspB40@OD2	LysB37@HD3	2.10	0.51
47	LeuB68@HB2	TrpB35@CE2	2.45	0.51
48	SerA61@O	LeuA55@HD12	2.12	0.50
49	PheB47@HD2	IleB33@HG13	1.77	0.50
50	ValB83@HG22	TrpB35@CZ3	2.45	0.50
51	**IleA82@HD11**	**LysB41@NZ**	**2.27**	**0.49**
52	ThrA83@HG23	IleA82@O	2.13	0.49
53	LeuA67@HD11	ValA13@CG2	2.43	0.48
54	AspA71@CG	ArgA52@HH22	2.16	0.48
55	AspB79@CG	SerB39@HG	2.17	0.48
56	GluB8@N	AlaB6@O	2.46	0.48
57	IleB85@HG13	IleB33@HA	1.96	0.48
58	ProA22@HD2	GlnA21@HG2	1.97	0.47
59	**GluA39@OE2**	**LysB51@CD**	**2.62**	**0.47**
60	**AspA84@HA**	**LysB41@HZ2**	**1.78**	**0.47**
61	GluB52@HG2	GluB52@O	2.15	0.47
62	TyrA10@HB2	LysA9@O	2.18	0.43
63	LeuA27@N	ValA26@CG1	2.80	0.43
64	LysB64@C	LeuB63@O	2.55	0.43
65	LeuB68@HD21	TyrB60@HB3	2.01	0.43
66	LysA24@HA	LysA24@HD3	1.90	0.42
67	SerA41@CB	AsnA40@O	2.67	0.42
68	SerA59@O	ValA17@HB	2.19	0.42
69	SerA61@N	SerA59@O	2.53	0.42
70	TyrA75@CZ	LysA30@HG3	2.54	0.42

71	SerA95@O	ValA8@HB	2.19	0.42
72	AsnB65@CB	LysB64@O	2.67	0.42
73	ValB83@HG22	TrpB35@CE3	2.55	0.42
74	ThrA14@N	ValA13@HG12	2.33	0.41
75	AsnA40@HD22	AsnA40@HA	1.60	0.41
76	SerA59@HA	SerA19@OG	2.19	0.41
77	SerA61@OG	HisA16@HA	2.20	0.41
78	**ProA80@HB3**	**LysB42@NZ**	**2.35**	**0.41**
79	GlyA93@CA	GlyA7@O	2.68	0.41
80	LysB41@N	ThrB38@O	2.54	0.41
81	HisB72@C	LysB71@O	2.59	0.41
82	TyrA75@CE1	LysA30@HG3	2.57	0.40
83	ThrB59@H	LysB57@C	2.25	0.40
84	ThrB59@N	LysB57@C	2.73	0.40
85	ValB83@N	LysB82@HG2	2.36	0.40

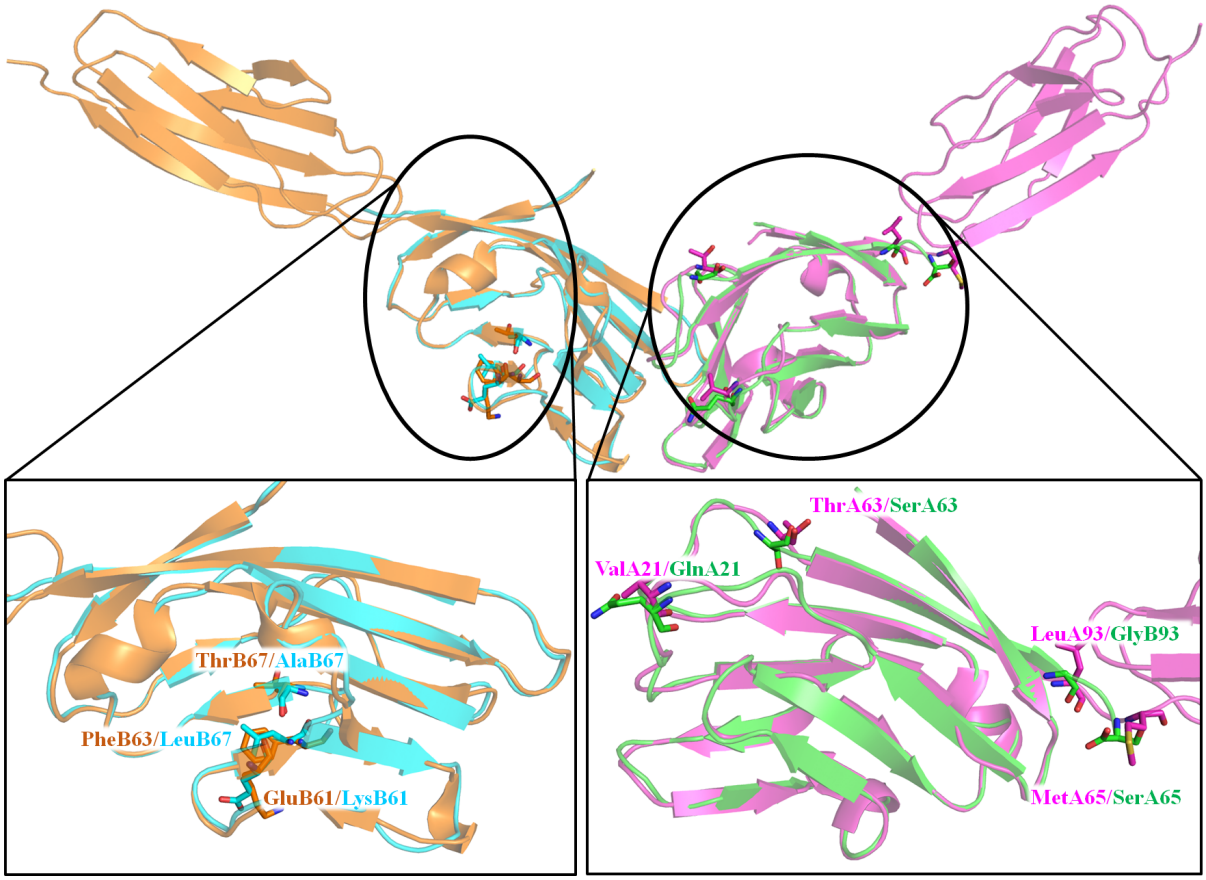

Figure S11. Structural overlap between *apo*-CD2 (orange, PDB ID: 1HNF), *apo*-CD58 (magenta, PDB ID: 1CCZ) and CD2-CD58 complex (CD2 in cyan and CD58 in green, PDB ID: 1QA9) to identify the residue differences. The mutated residues in the crystal structure of CD2-CD58 and the corresponding residues in the *apo* forms are shown as sticks.

MD analysis of CD2-CD58

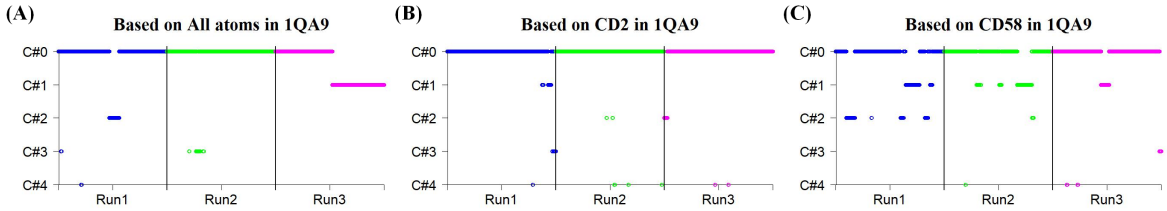

Figure S12. Combined clustering analysis of the triplicate MD trajectories for the CD2-CD58 complex by considering the RMSD of all heavy atoms (A) in the complex, (B) in CD2, and (C) in CD58.

Table S6. All-atom root mean square deviation (RMSD) values between the centroids of various clusters for the combined clustering analysis of the three replicate MD simulations, shown in Figure S4. The RMSD value (in Å) is indicated for clustering based on all atoms in the complex (for CD58 based clustering/for CD2 based clustering) for each cluster pair.

Cluster	C#0	C#1	C#2	C#3	C#4
C#0	0	1.74 (1.15/1.88)	1.78 (0.85/1.01)	1.52 (1.96/1.51)	2.19 (0.90/0.81)
C#1		0	2.16 (1.57/1.77)	2.80 (1.76/1.15)	2.78 (1.19/1.91)
C#2			0	1.24 (2.25/1.47)	2.77 (0.99/1.09)
C#3				0	2.69 (1.88/1.40)
C#4					0

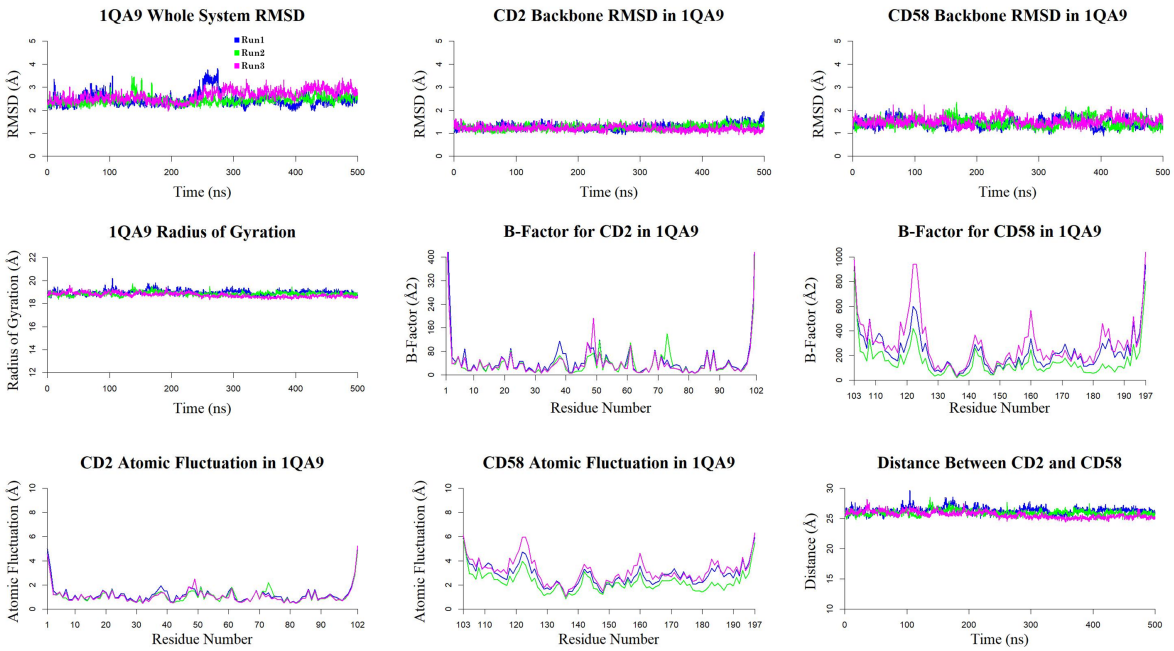

Figure S13. System stability analysis for the triplicate MD simulations for the CD2-CD58 complex.

MD analysis of apo-CD2 and apo-CD58

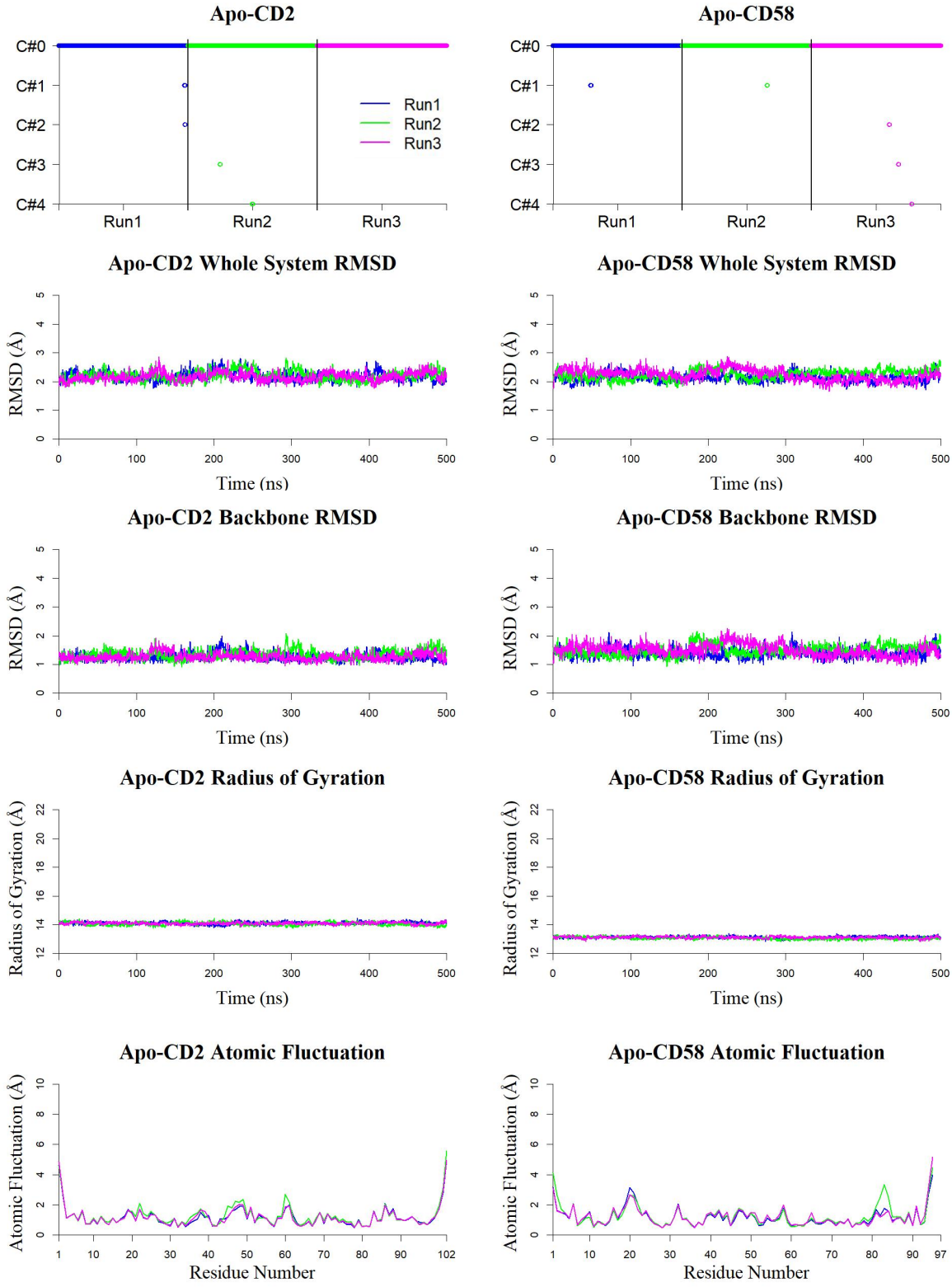

Figure S14. System stability analysis from the triplicate MD simulations for *apo*-CD2 and *apo*-CD58.

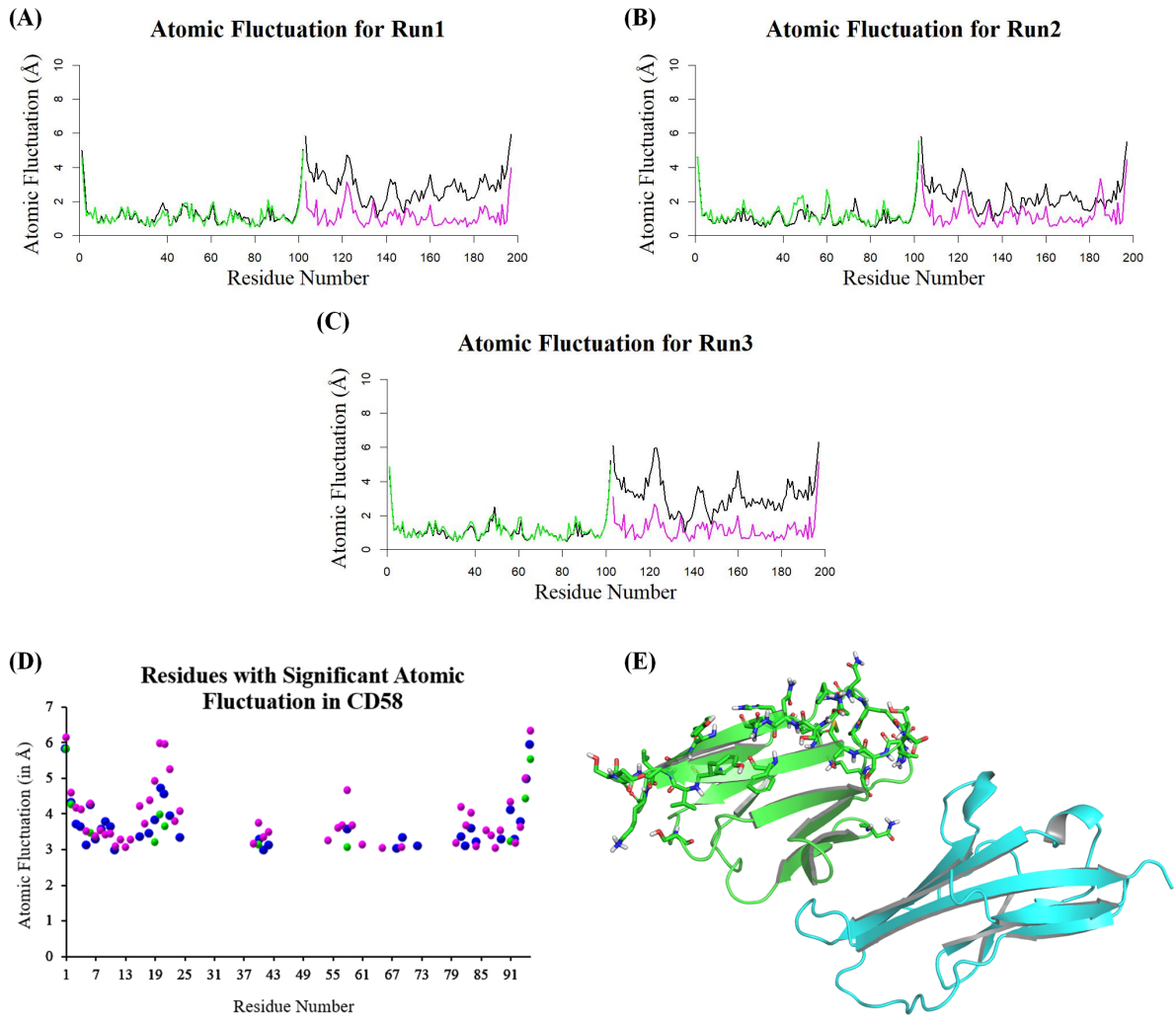

Figure S15. Comparative atomic fluctuation analysis (A), (B), and (C) between *apo*-CD2 (green) and *apo*-CD58 (magenta) with the complex (black) from the triplicate MD simulations, (D) residues numbers which exhibited atomic fluctuation higher than 3 Å in the CD58 during the triplicate MD simulations, and (E) topographical location of the residues exhibiting atomic fluctuation higher than 3 Å during simulations.

MM/GBSA ΔG_{bind} analysis

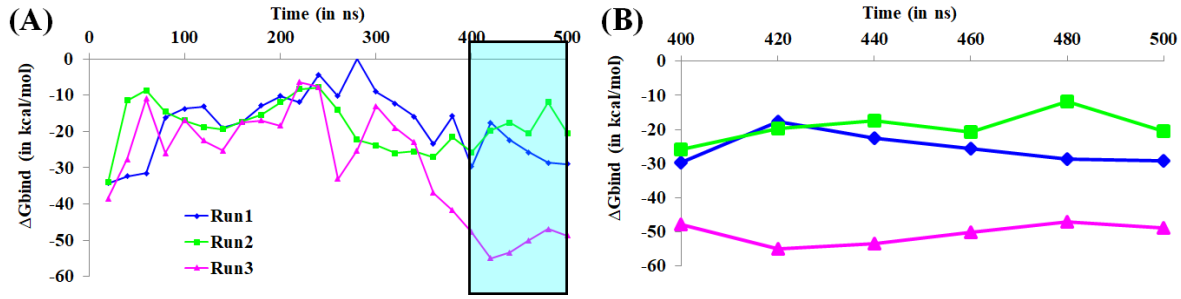

Figure S16. MM/GBSA ΔG_{bind} calculation for each of the 20 ns simulation trajectory for the CD2-CD58 complex (A) over 500 ns, (B) from 400 ns to 500 ns.

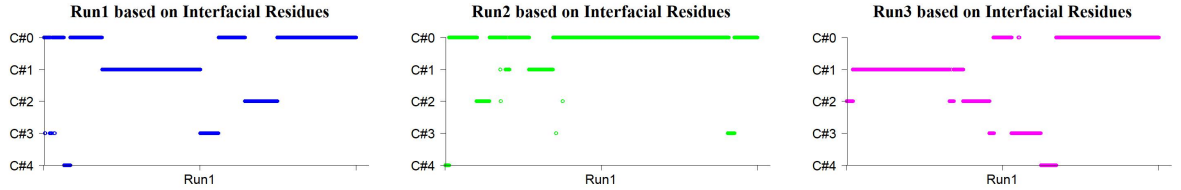

Figure S17. Clustering analysis of selected interfacial residues in the CD2-CD58 complex during Run1, Run2, and Run3, performed for each individual simulation replicate.

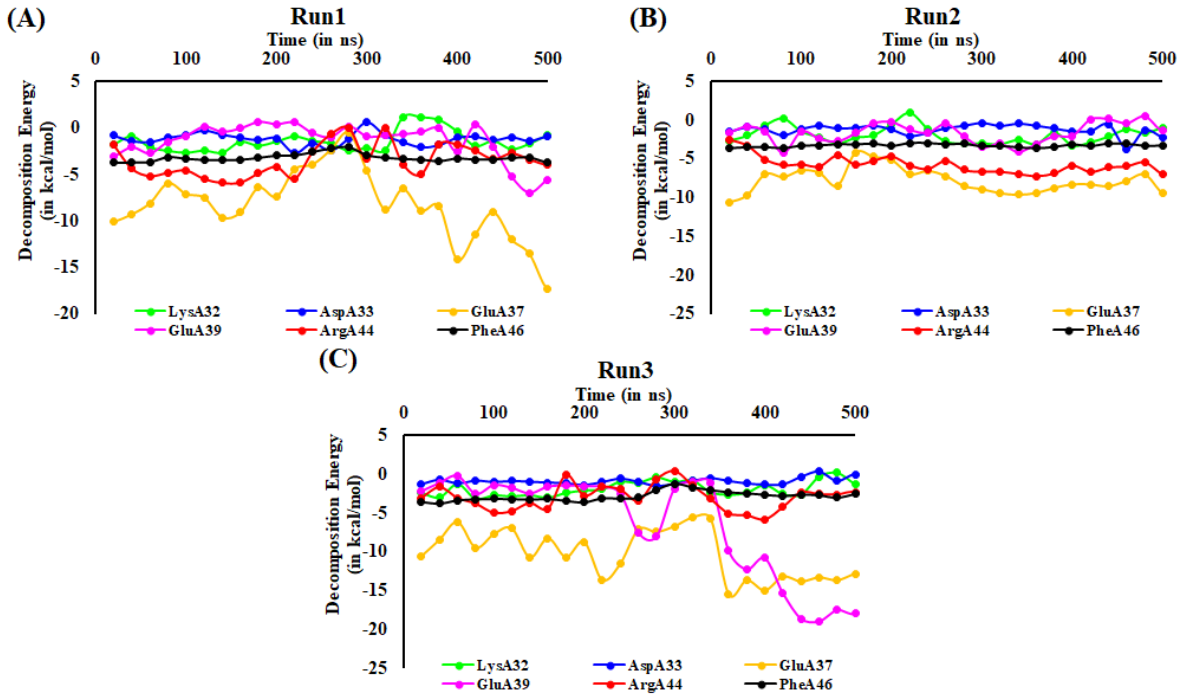

Figure S18. Per-residue decomposition energy analysis for the selected CD58 residues over the simulation period during (A) Run1, (B) Run2, and (C) Run3.

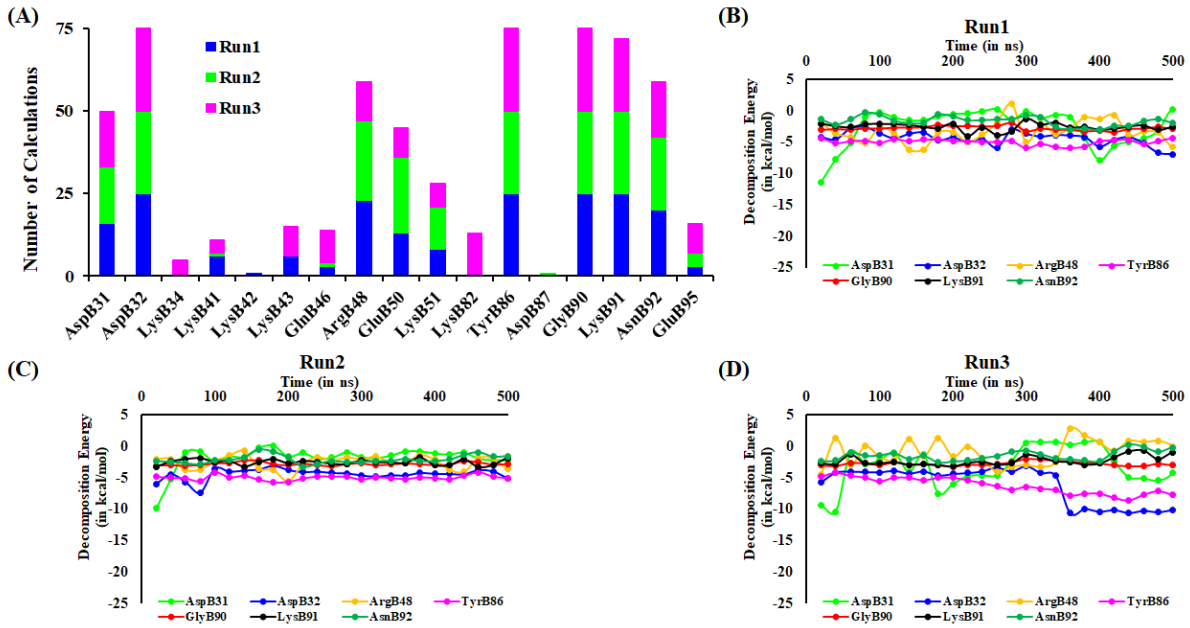

Figure S19. Identification of the important CD2 residues. (A) Occupancy analysis for the residues exhibiting decomposition energy lower than -1.00 kcal/mol for the interaction with CD58 during the replicate MD simulations. Per-residue decomposition energy analysis of the selected CD58 residues over the simulation period for (B) Run1, (C) Run2, and (D) Run3.

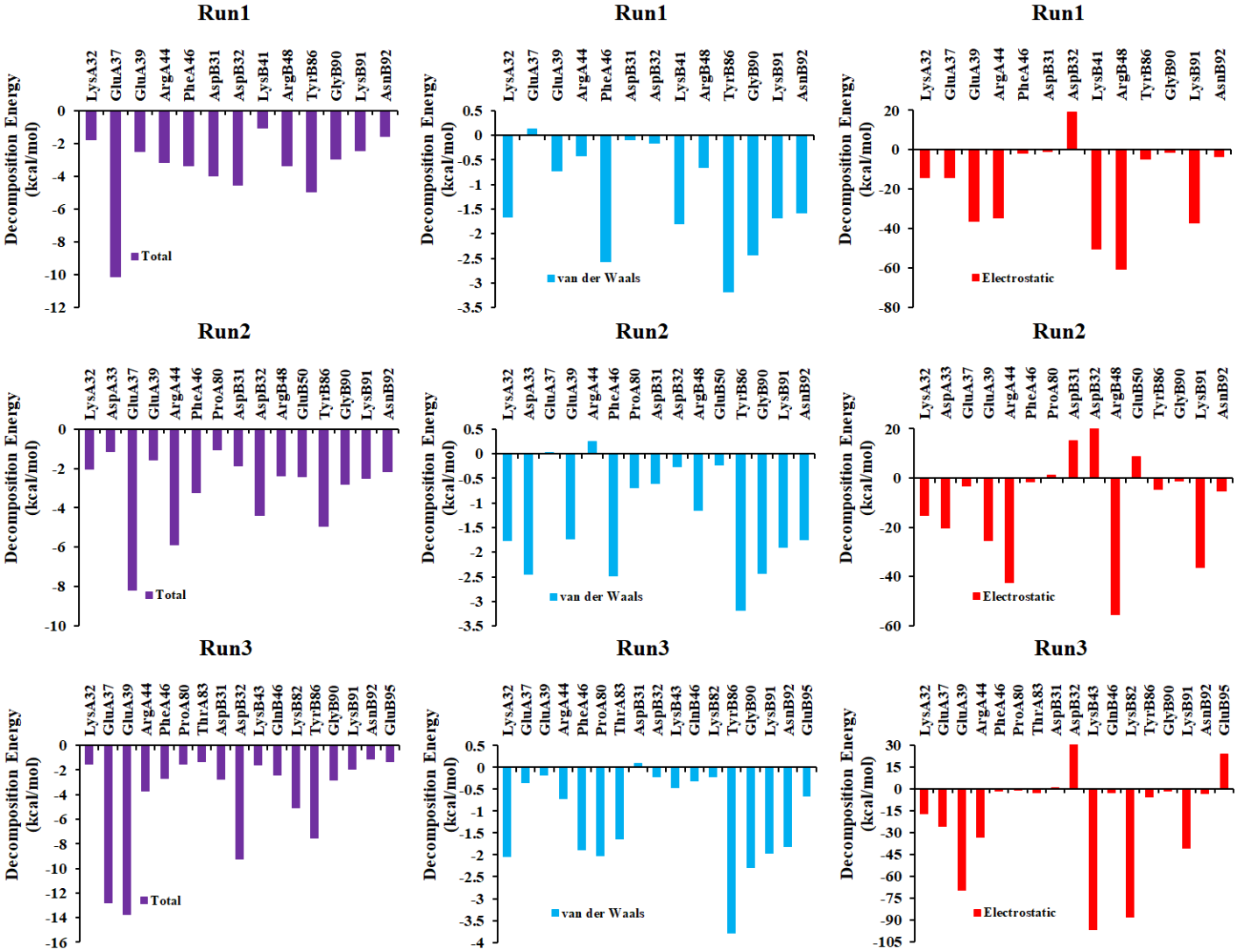

Figure S20. Per-residue decomposition energy analysis for the important residues identified from the binding energy calculations on the most populated cluster in Run1, Run2, and Run3. The van der Waals and electrostatic contributions from these residues are also provided.

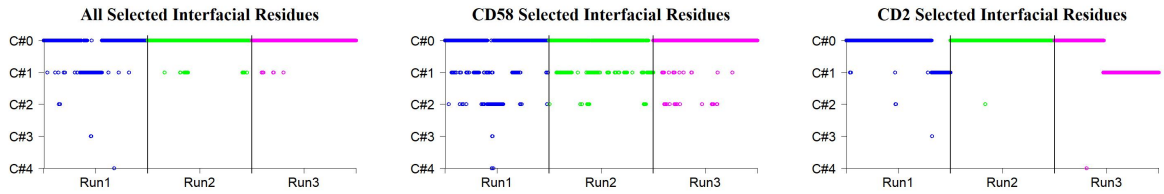

Figure S21. Combined clustering analysis of the selected interfacial residues in the CD2-CD58 complex during the replicate MD simulations Run1, Run2, and Run3.

Virtual screening protocol

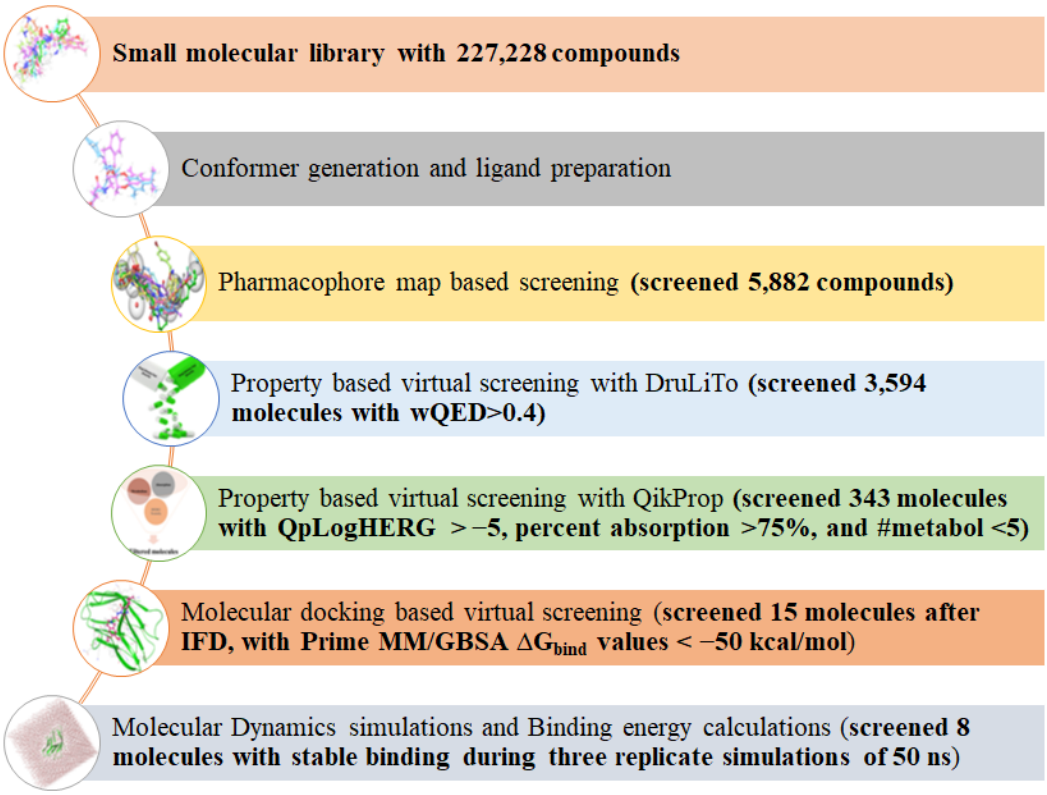

Figure S22. Virtual screening protocol followed for the identification of the CD58-binding inhibitors of the CD2-CD58 interactions.

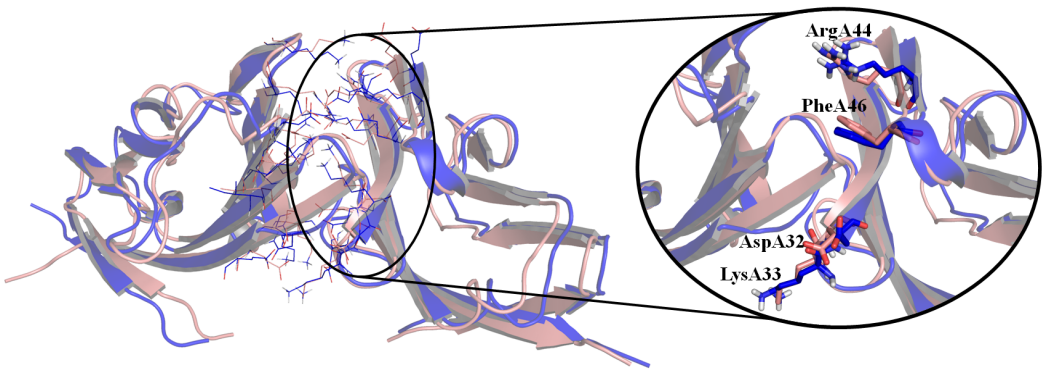

Figure S23. Comparative analysis between the last frame (peach) and the centroid of the most populated cluster (blue) (RMSD 1.43 Å). Residues important for the interfacial

interactions are shown as lines and those considered for the pharmacophore map development from the crystal structure are shown as sticks.

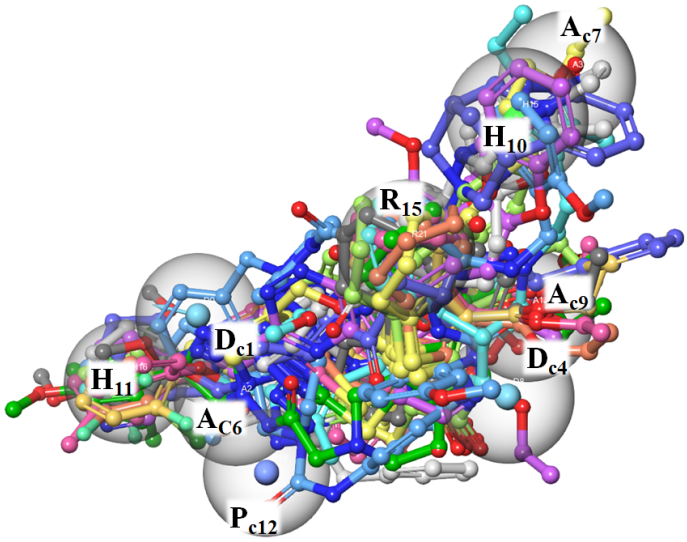

Figure S24. Top twenty compounds (based on PhaseScreen score) screened after the pharmacophore-based virtual screening, overlapped with generated pharmacophore shown in Figure 4C.

Properties based filtering

The 5,882 compounds obtained after the pharmacophore-based screening were submitted to DruLiTo based evaluation of drug-likeness (Table S7). 3,594 compounds, with wQED larger than 0.4, were selected for the next screening step (Figure S14), in which additional pharmacokinetic parameters were evaluated using QikProp. The filtering criteria for the QikProp based filtering were QpLogHERG > -5, percent absorption >75%, and #metabol < 5. The various molecular properties, which are reported to correlate with the pharmacokinetic behavior of small molecules were within the recommended ranges (Figure S17-S18).

Table S7. Number of molecules filtered after applying the various drug-likeness rules evaluated using DruLiTo.

Drug likeness rule	Number of molecules filtered
BBB Rule	236
CMC-50 like Rule	170
Ghose Filter	1,921
Lipinski's Rule	3,572
MDDR-like Rule	5,408
Veber Rule	2,686
uwQED	3,328
wQED	3,594

All Filters	28

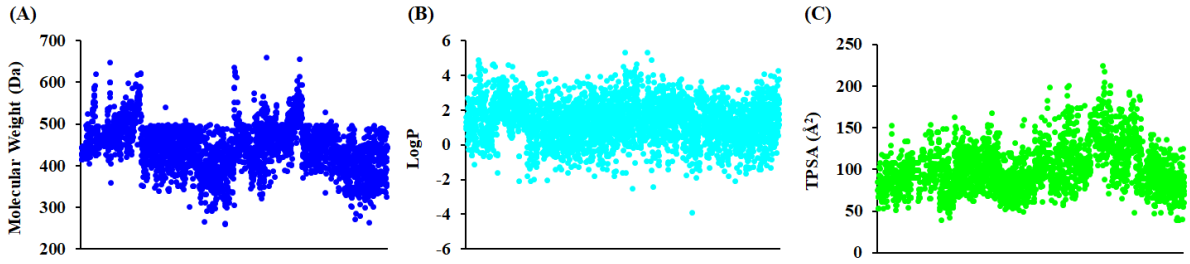

Figure S25. Molecular property distribution of (A) molecular weight, (B) LogP, and (C) total polar surface area (TPSA) of the molecules with wQED > 0.4. The molecular weight distribution > 90% of the selected compounds was between 350 to 500 Da, whereas the LogP distribution range was 1 to 4, indicating a potential for membrane permeation. The total polar surface area for the majority of the selected molecules was below 140 Å^2, supporting their potential to cross the membrane barrier.

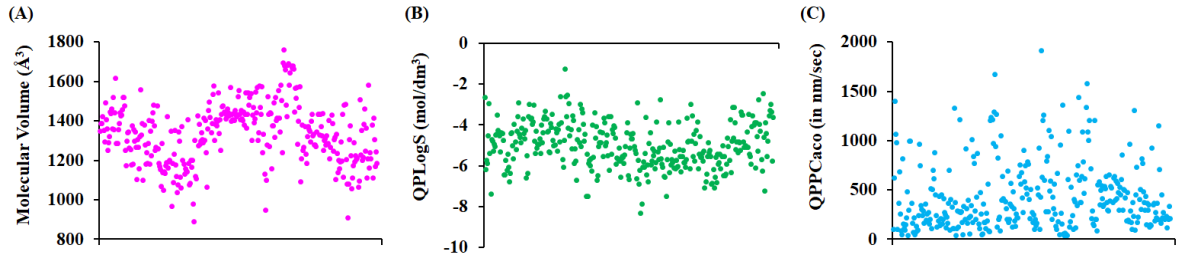

Figure S26. Factors correlated to the pharmacokinetic properties, *i.e.*, (A) molecular volume, (B) Predicted aqueous solubility (QPLogS), and (C) Predicted apparent Caco-2 cell permeability (QPPCaco) for the molecules with QpLogHERG > −5, percent absorption >75%, and #metabol <5, filtered after QikProp based property calculations. The number of compounds that exhibited the selected features in the allowed range was 343. These filtered compounds exhibited a molecular volume in the range of 800 to 1,800 Å^3 (recommended range: 500–2,000 Å^3), predicted aqueous solubility (QPLogS) in the range −8 to −2 mol/dm^3 (recommended range –6.5 to 0.5 mol/dm^3), and predicted apparent Caco-2 cell permeability (QPPCaco) >50 nm/sec (< 25 is considered as poor as per the recommended range).

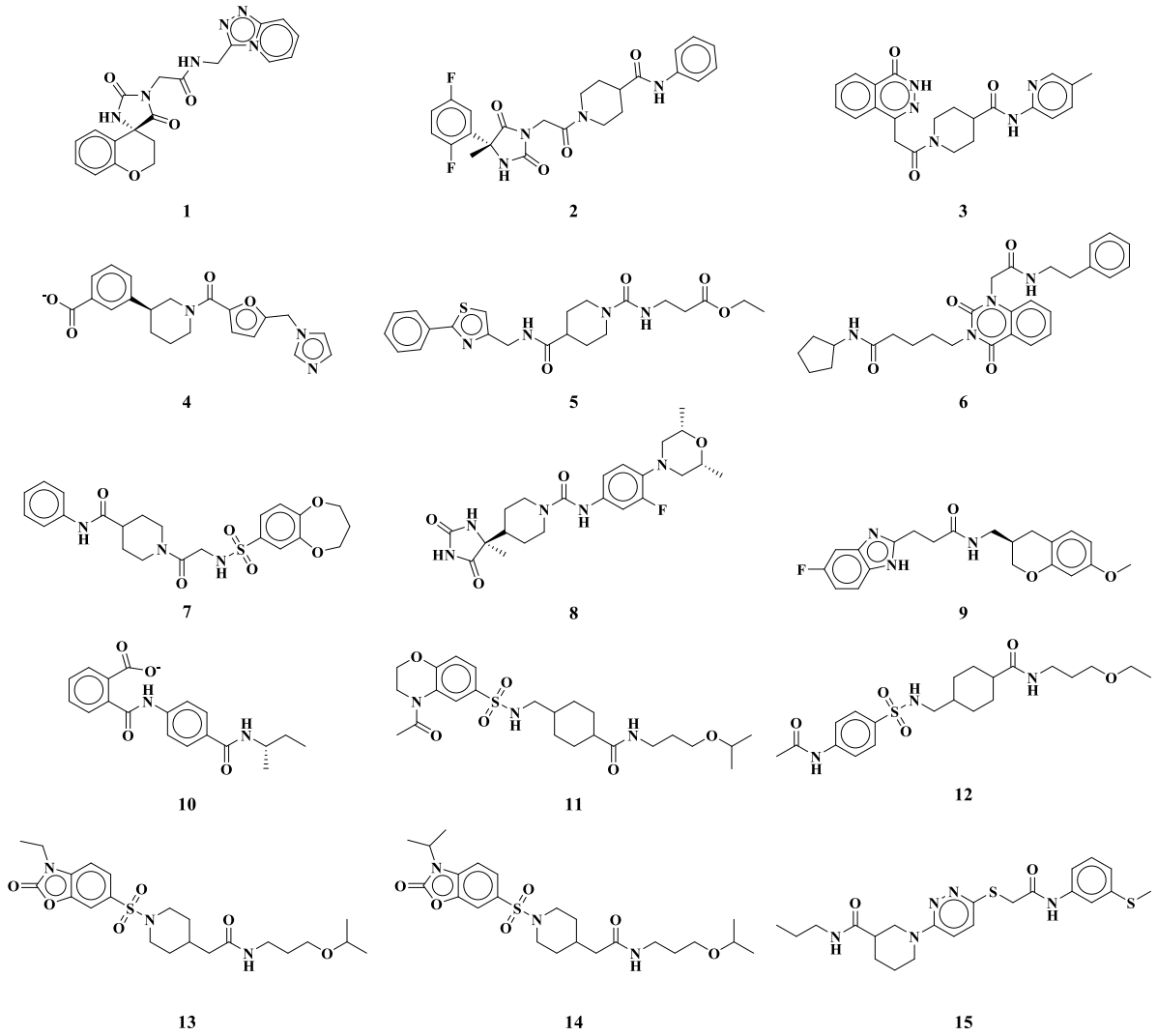

Figure S27. 2D structure of the compounds selected after molecular docking based virtual screening, which than were further submitted to MD simulations to evaluate their binding to CD58.

Promising hits analysis

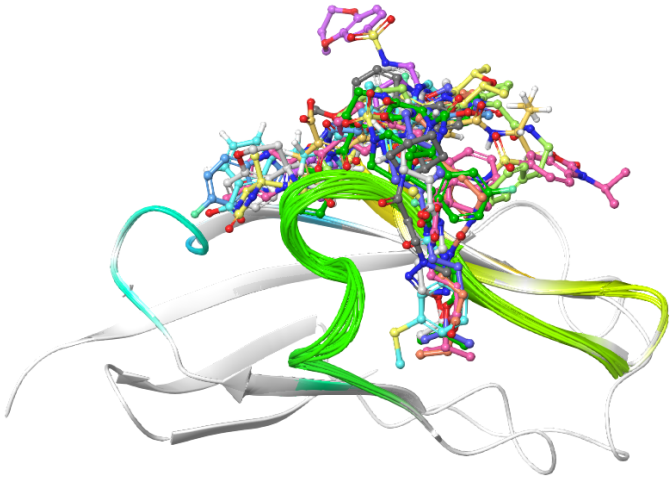

Figure S28. Structural overlap of the top-scoring poses for the selected molecules (**compounds 1-15**). The structural components of CD58 which were not considered for minimization during IFD are shown as grey.

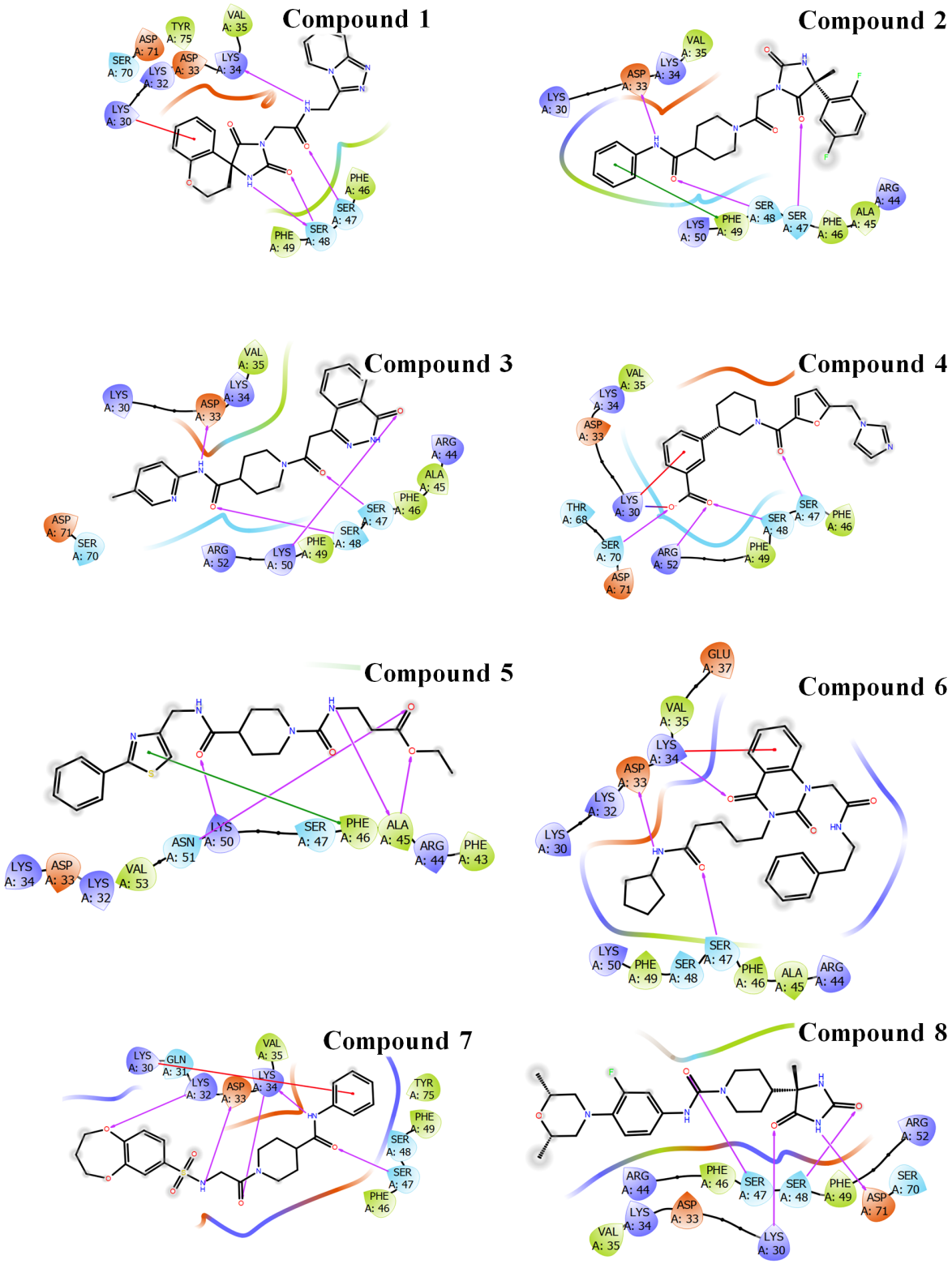

Figure S29. 2D interactions diagram for **compounds 1-8**. Hydrogen bonds: magenta lines; π···π interactions: green lines; NH···π/ CH···π interactions: red lines; salt bridges: blue and red gradient filled lines.

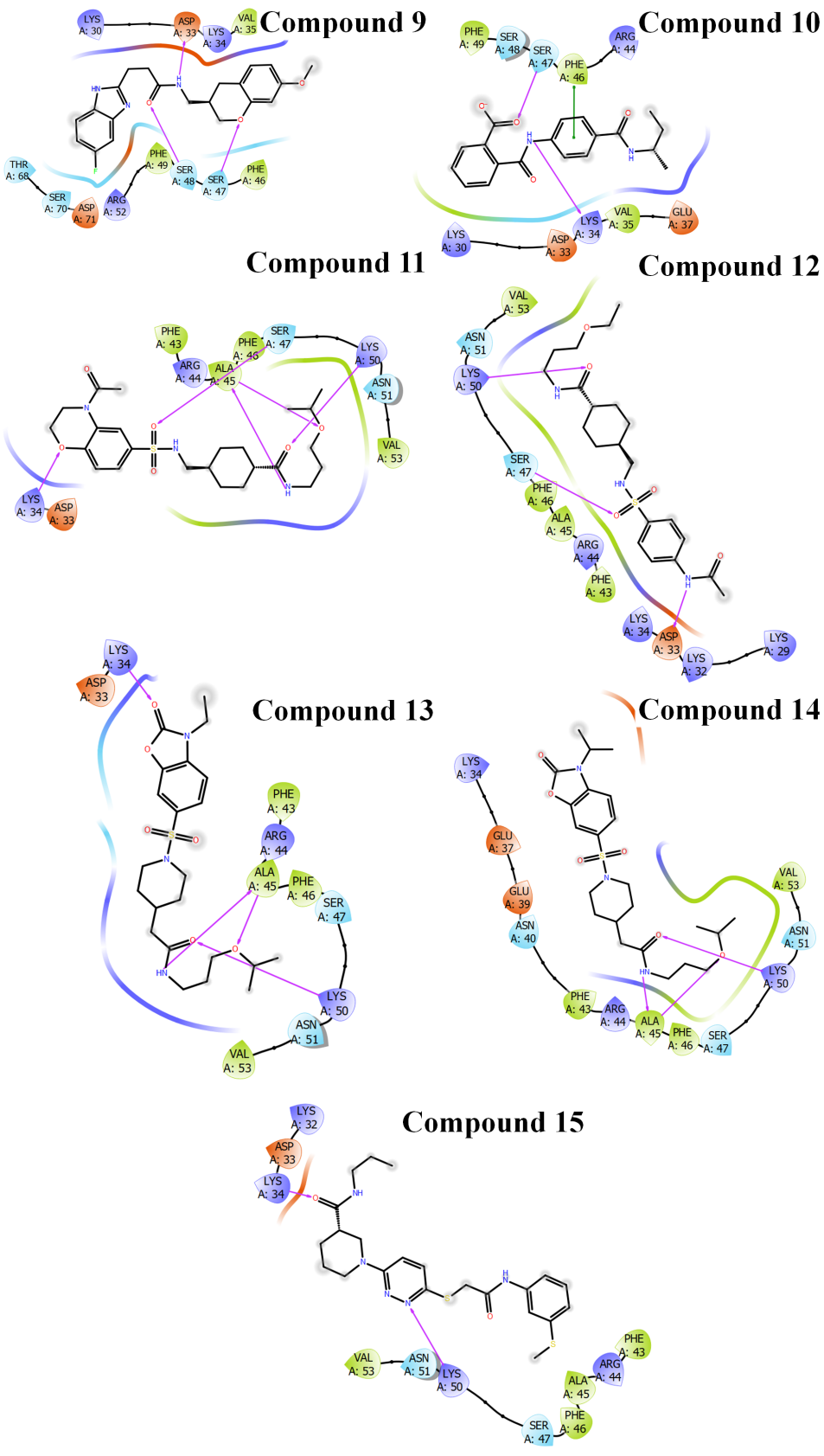

Figure S30. 2D interactions diagram for **compounds 9-15**. Hydrogen bonds: magenta lines; π···π interactions: green lines; NH···π/ CH···π interactions: red lines; salt bridges: blue and red gradient filled lines.

Figure S31. Structural superimposition of top three poses for **compounds 1-15**, based on MM/GBSA ΔG_{bind} values, after Induced Fit Docking (IFD). CD58 is shown as cartoon representation and the docked ligands as stick representation (Pose1: blue, Pose2: green, Pose3: magenta).

Table S8. Prime-MM/GBSA ΔG_{bind} (in kcal/mol) values for the top three poses obtained after the Induced Fit Docking (IFD) of compounds **1-15**.

Molecule	Pose1	Pose2	Pose3
Compound 1	−65	−50	−46
Compound 2	−60	−52	−51
Compound 3	−59	−58	−55
Compound 4	−58	−57	−56
Compound 5	−56	−52	−51
Compound 6	−59	−58	−56
Compound 7	−53	−52	−51
Compound 8	−54	−52	−50
Compound 9	−54	−44	−43
Compound 10	−51	−50	−50
Compound 11	−57	−53	−52
Compound 12	−51	−50	−50
Compound 13	−50	−49	−46
Compound 14	−53	−53	−51
Compound 15	−54	−54	−52

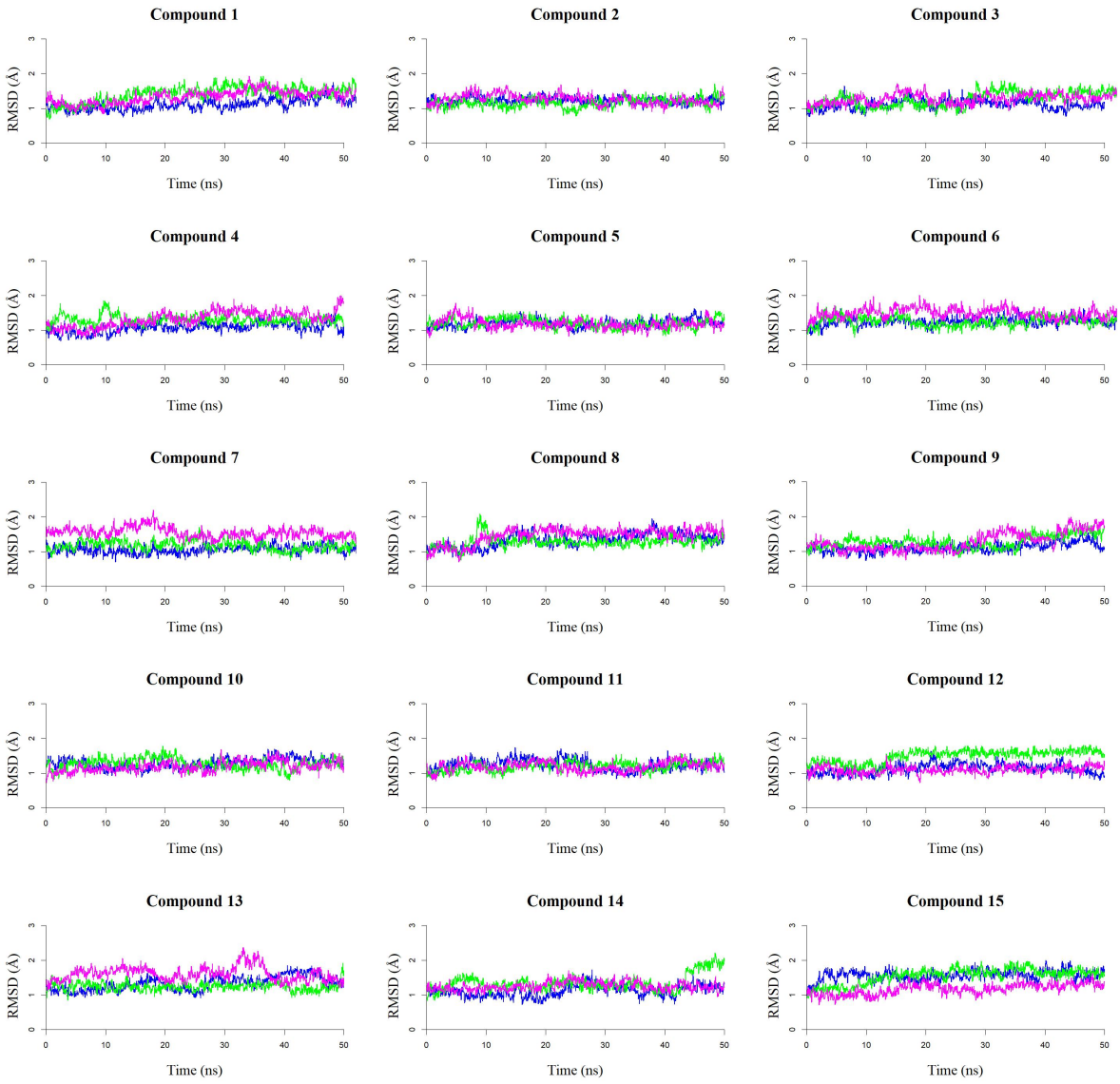

Figure S32. Analysis of the protein backbone RMSD for the MD simulations of compounds **1-15** in the three binding conformations (Pose1 in blue, Pose2 in green, and Pose3 in magenta).

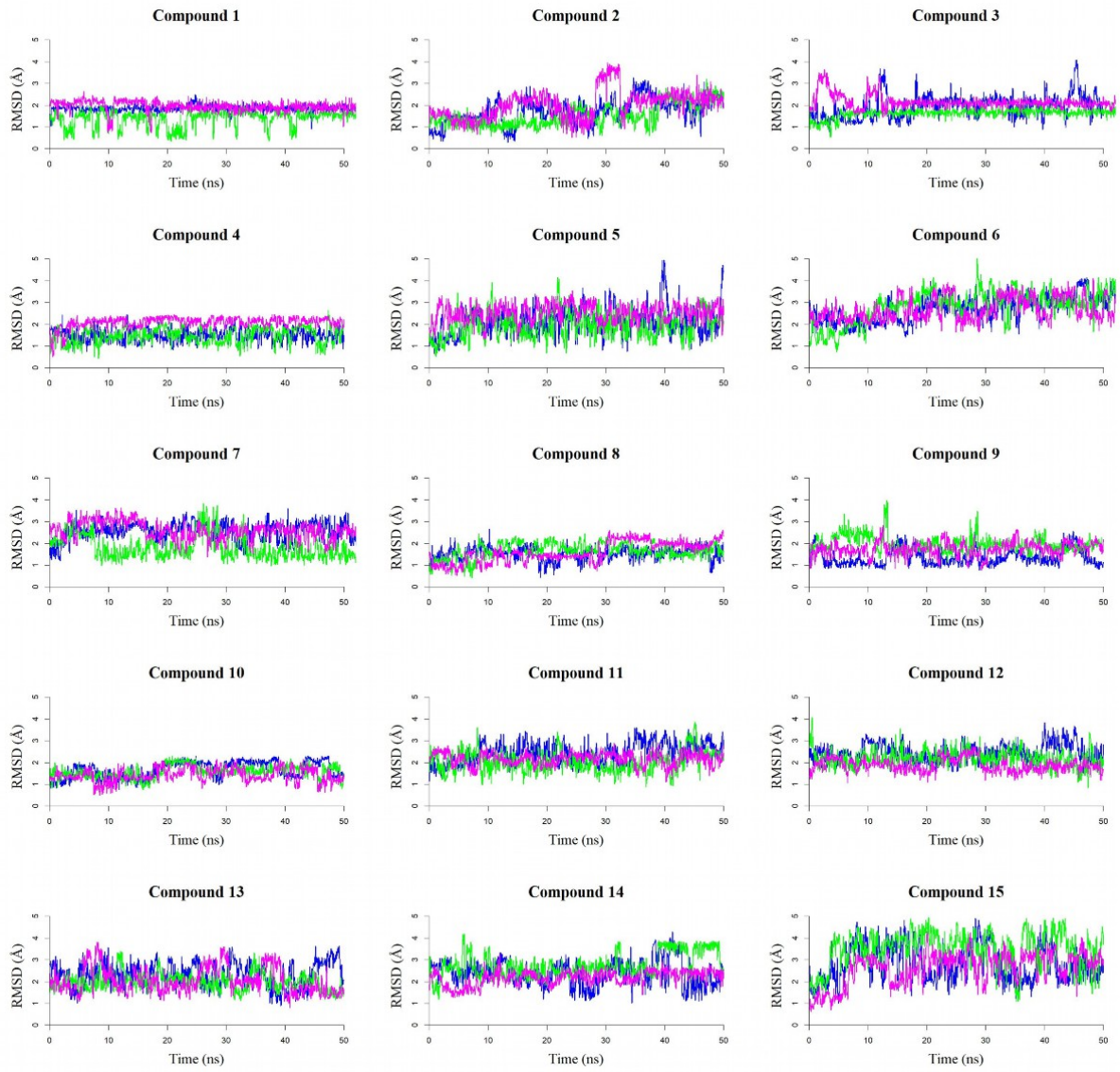

Figure S33. Analysis of the ligand RMSD for the MD simulations of compounds **1-15** in the three binding conformations (Pose1 in blue, Pose2 in green, and Pose3 in magenta).

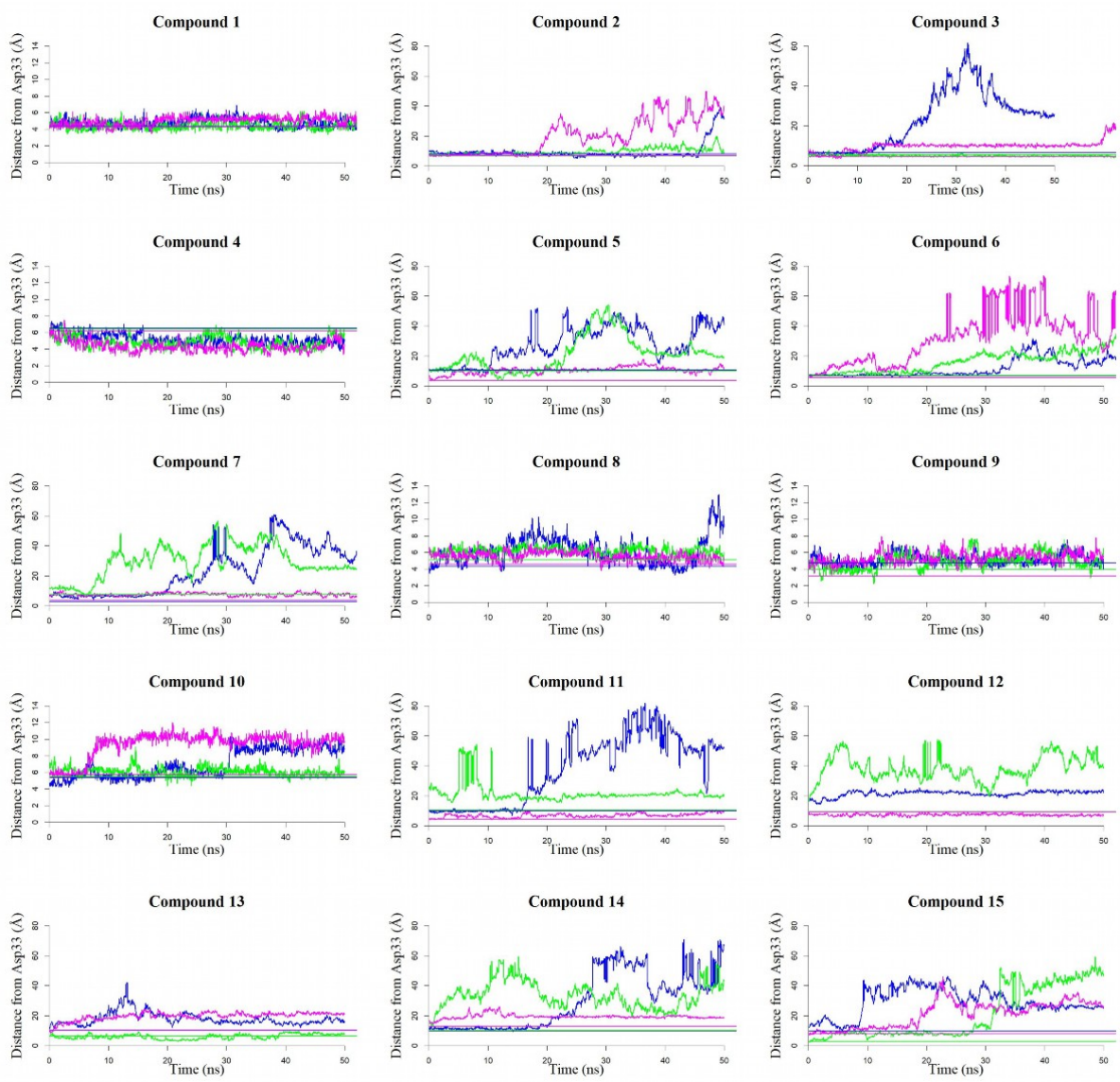

Figure S34. Analysis of the distance between the centroid of the bound ligand and AspA33 (residue located at the CD2 binding site) for the MD simulations of compounds **1-15** in the binding conformations (Pose1 in blue, Pose2 in green, and Pose3 in magenta). The initial distances are shown as straight lines in the respective colors for each Pose.

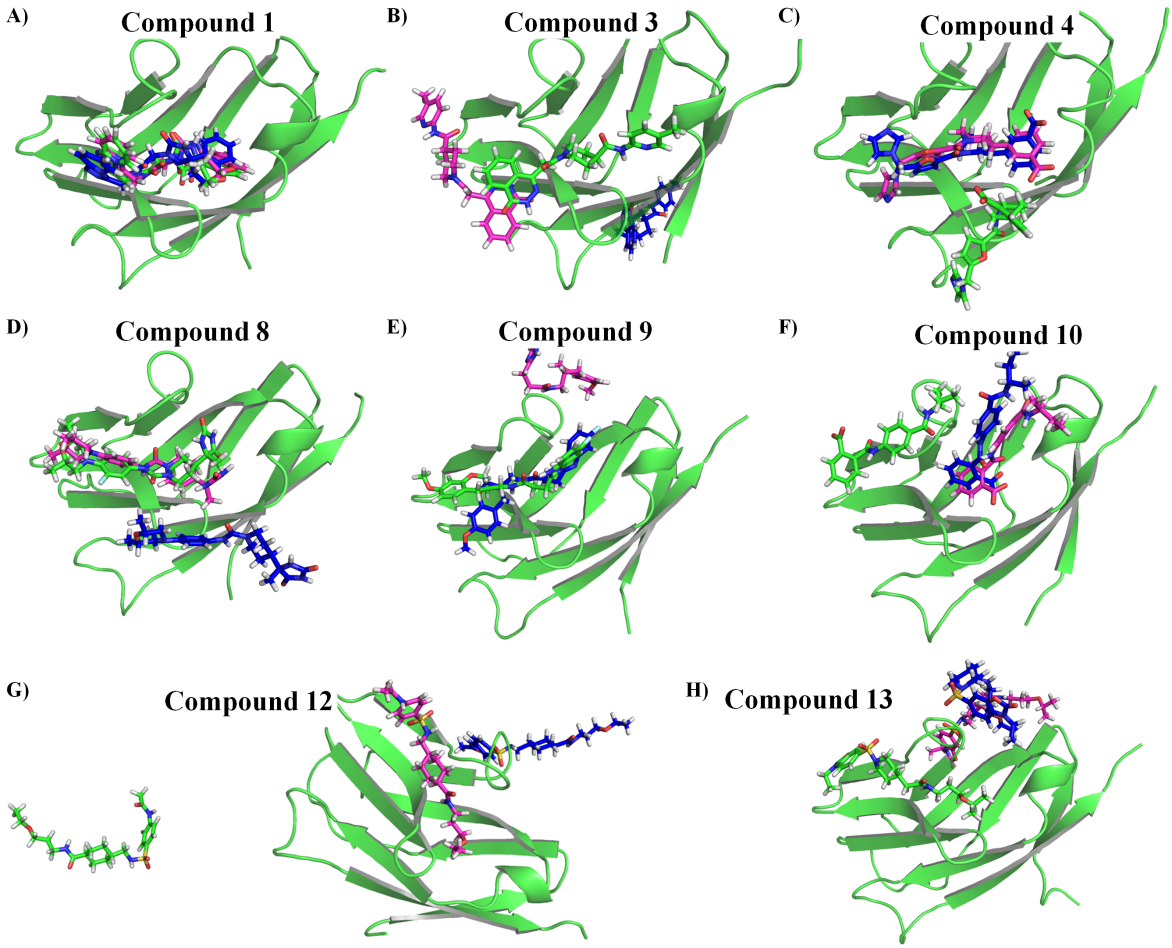

Figure S35. Superimposition of the ligand position at the end of MD simulations for the eight selected ligands (Pose1 in blue, Pose2 in green, and Pose3 in magenta). CD58 is shown as a cartoon representation.

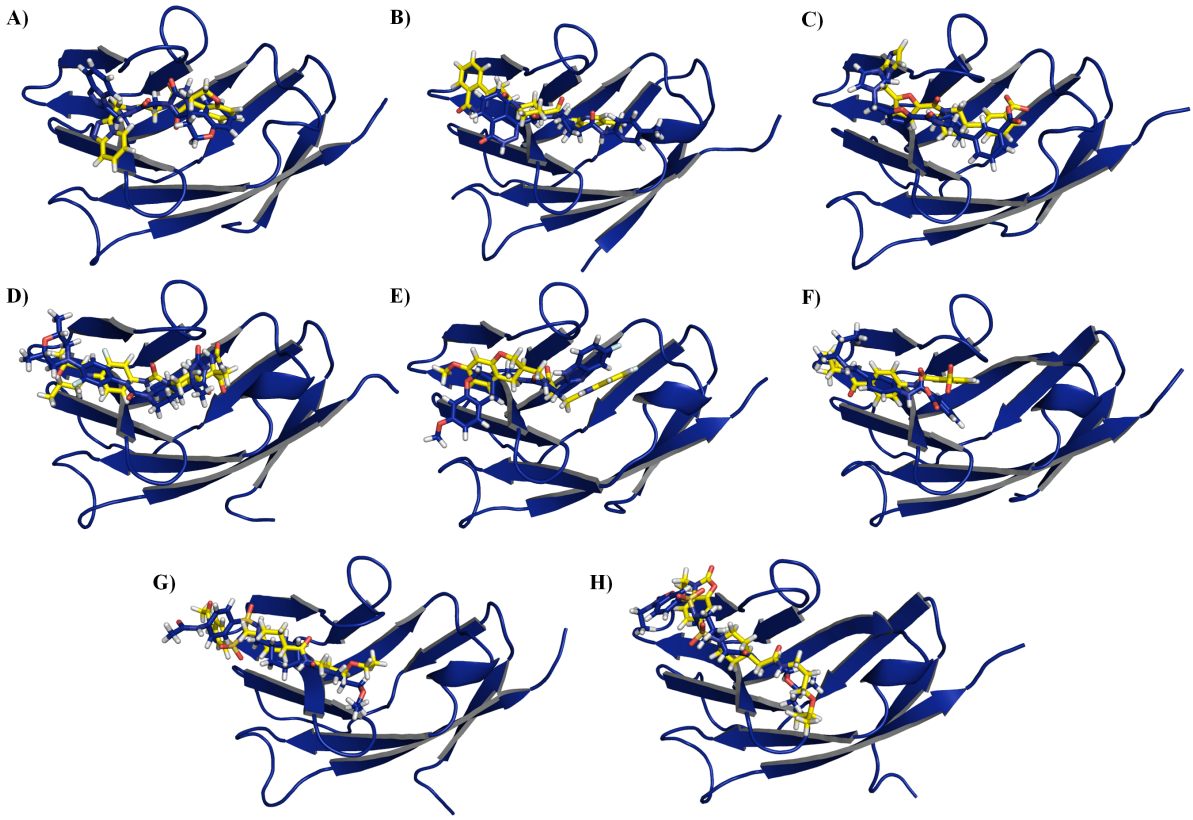

Figure S36. Overlap between the docked (yellow) and stable conformation after MD simulations (blue) for (A) **compound 1**, (B) **3**, (C) **4**, (D) **8**, (E) **9**, (F) **10**, (G) **12**, and (H) **13**. CD58 is shown as cartoon representation.

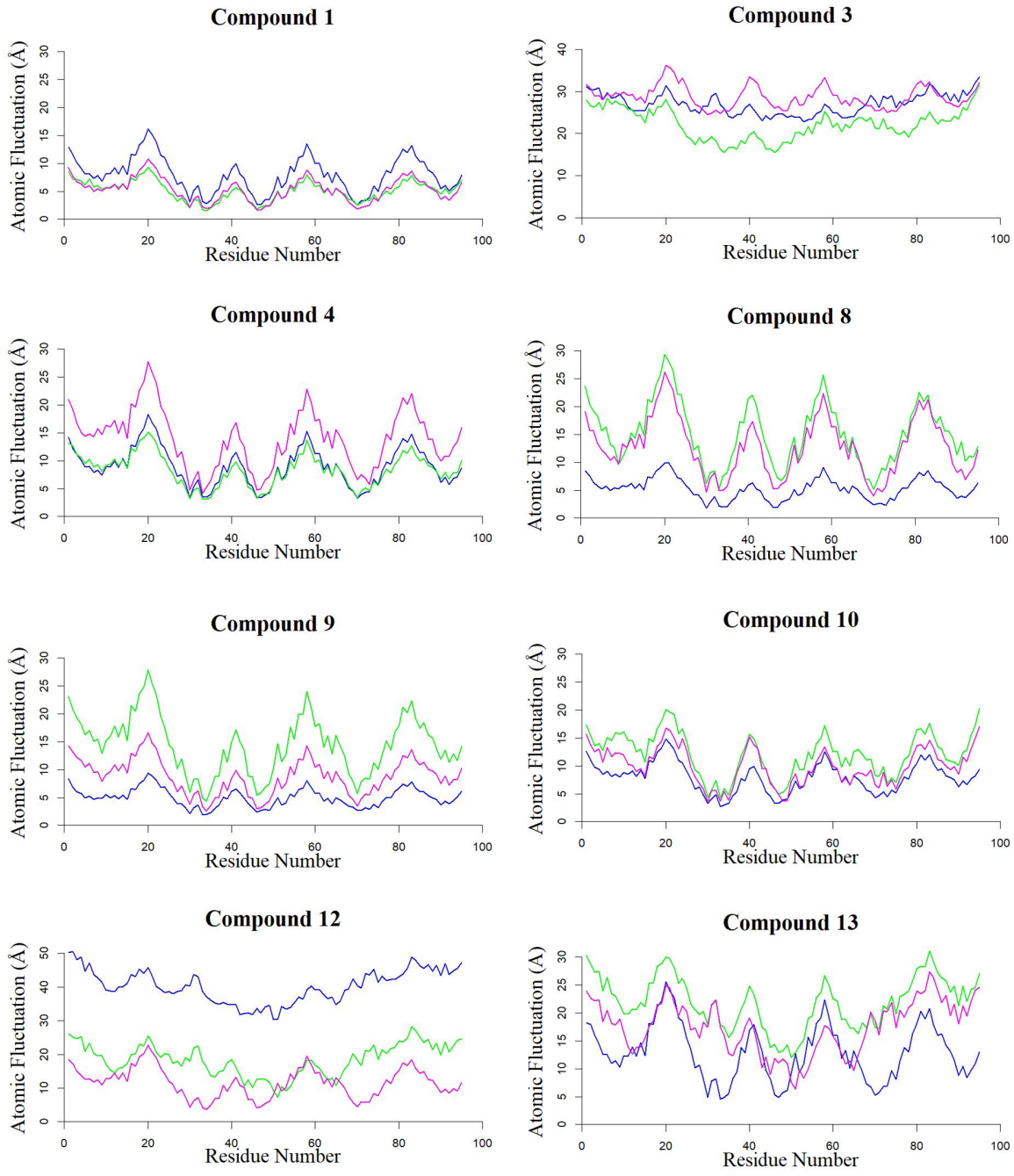

Figure S37. Atomic fluctuation analysis for the selected eight ligands (Pose1 in blue, Pose2 in green, and Pose3 in magenta).

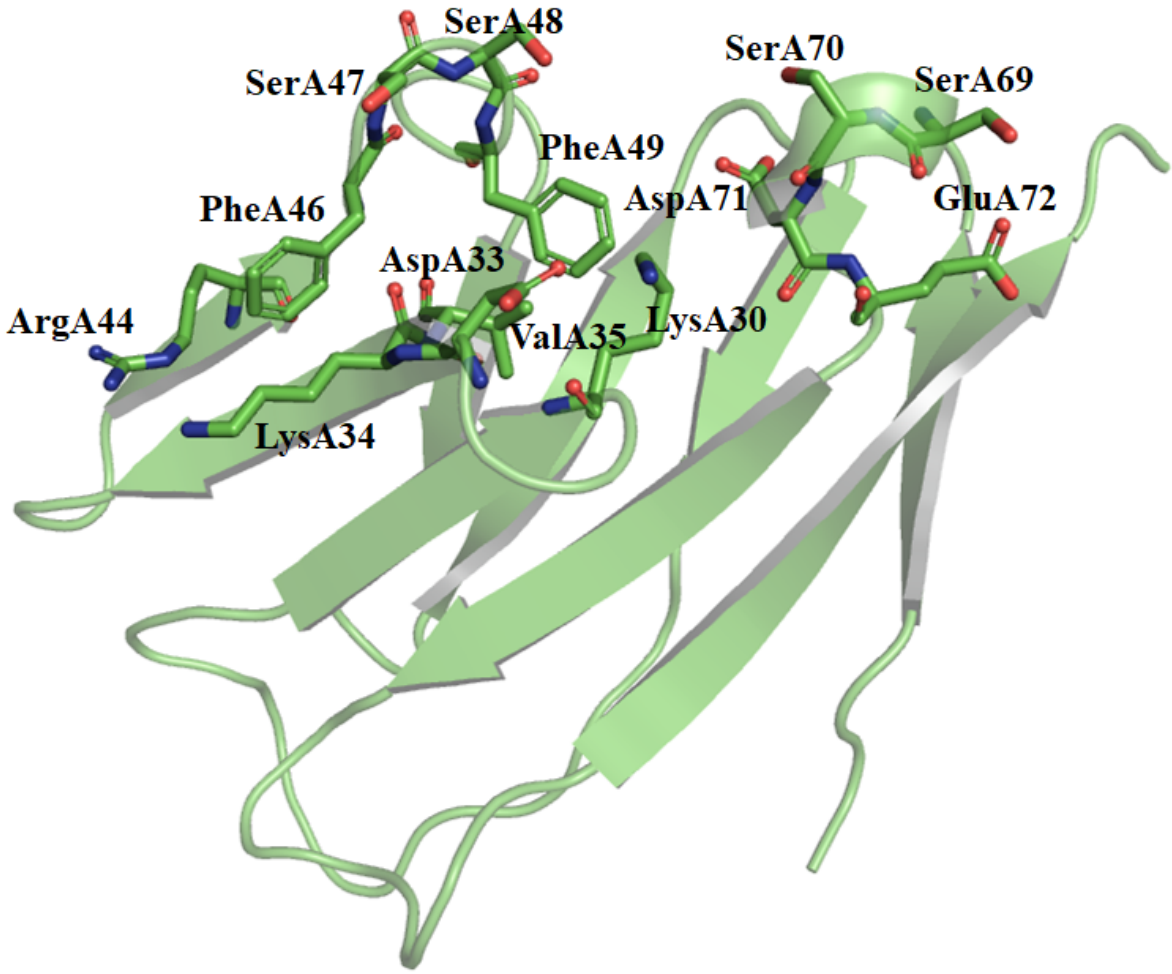

Figure S38. CD58 residues exhibiting the lowest atomic fluctuation (< 2 Å) for the selected ligands during the MD simulations. CD58 is shown as cartoon representation.

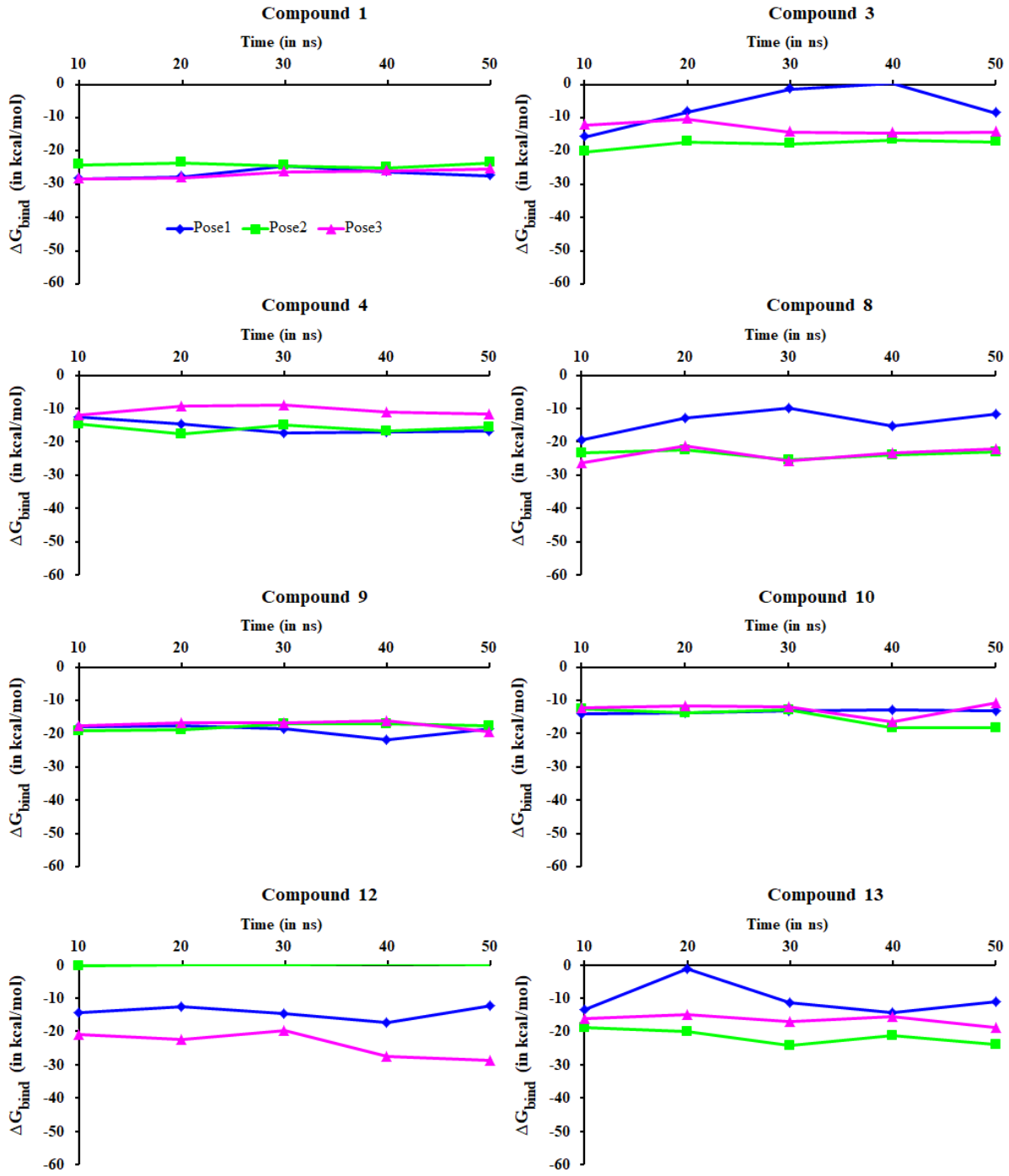

Figure S39. MM/GBSA ΔG_{bind} values (in kcal/mol) of the three different binding conformations (Pose1 in blue, Pose2 in green, and Pose3 in magenta) for the selected eight ligands over the 50 ns MD simulations.

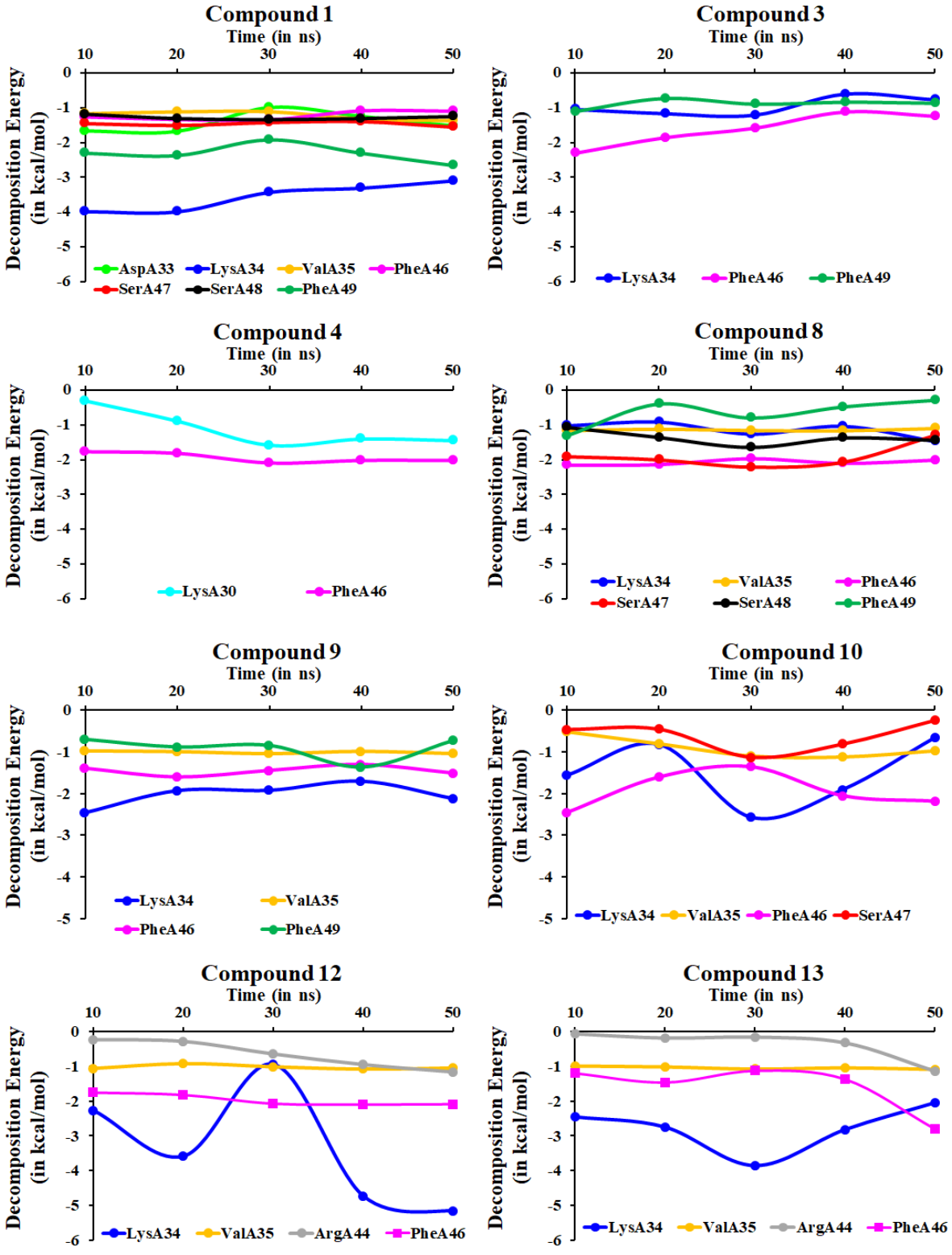

Figure S40. Per-residue decomposition energy analysis for the most stable binding conformation of the selected eight ligands over the 50 ns MD simulation.

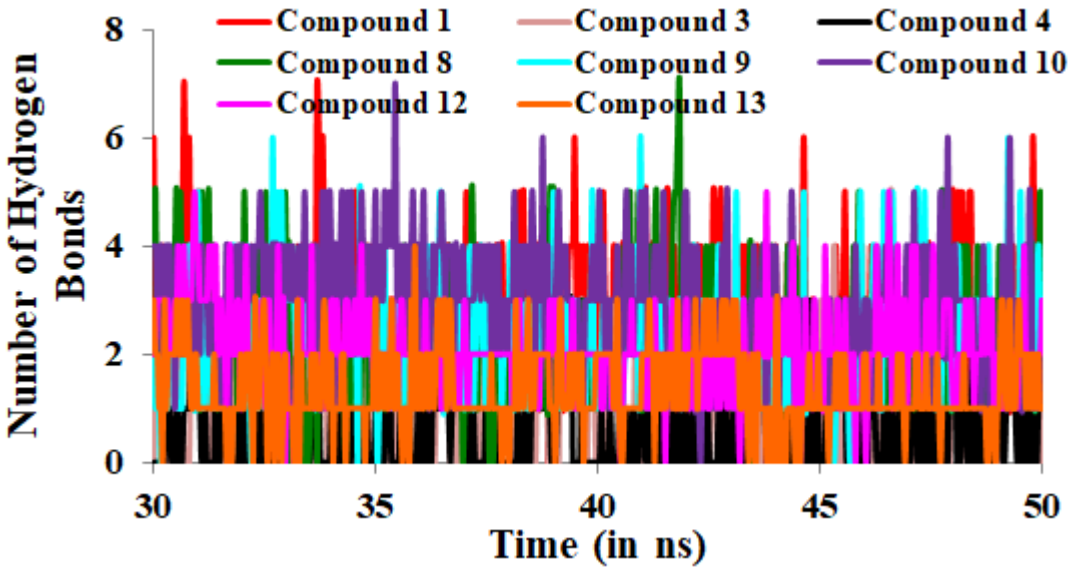

Figure S41. Number of hydrogen bonds between CD58 and the most stably bound binding conformation of the selected compounds over the last 20 ns MD simulation.

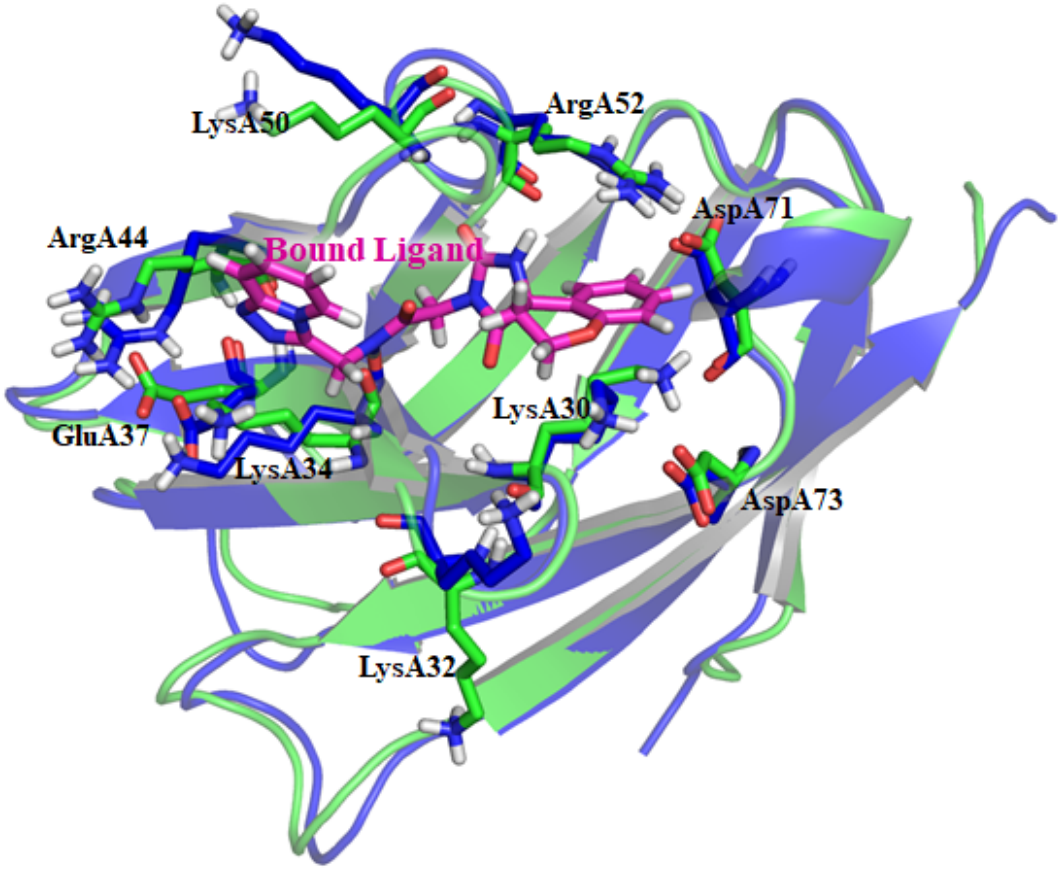

Figure S42. Comparative analysis between the initial (blue) and final (green) structure of CD58 in complex with ligands. The ionization state of the amino acids was maintained the same as that predicted using PROPKA (Figure S1). The ionizable residues which were close to the ligand are shown as sticks.

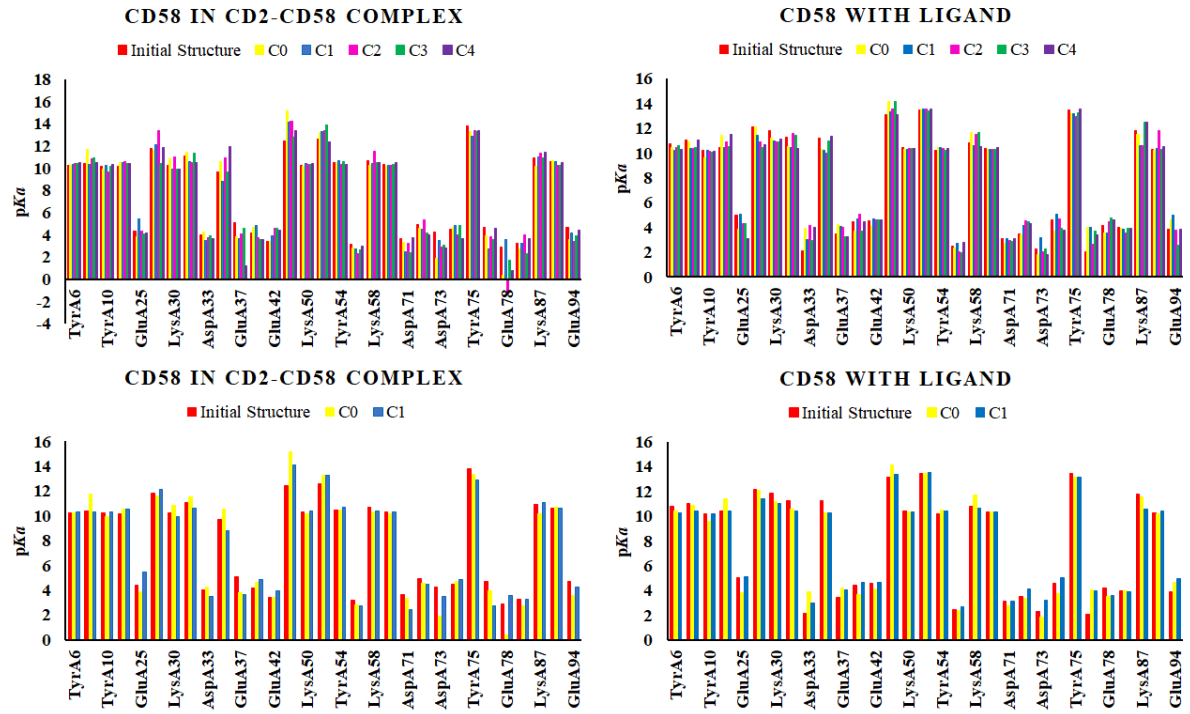

Figure S43. PROPKA analysis (http://server.poissonboltzmann.org/pdb2pqr)[12] on CD58 (centroid structure) in the various clusters after MD simulations of CD2-CD58 complex and complex with the ligand. For clarity, only the top two clusters (ranked by population size) is shown in the lower panel.

Comparative analysis between results from GAFF/FF99SB and GAFF2/FF14SB

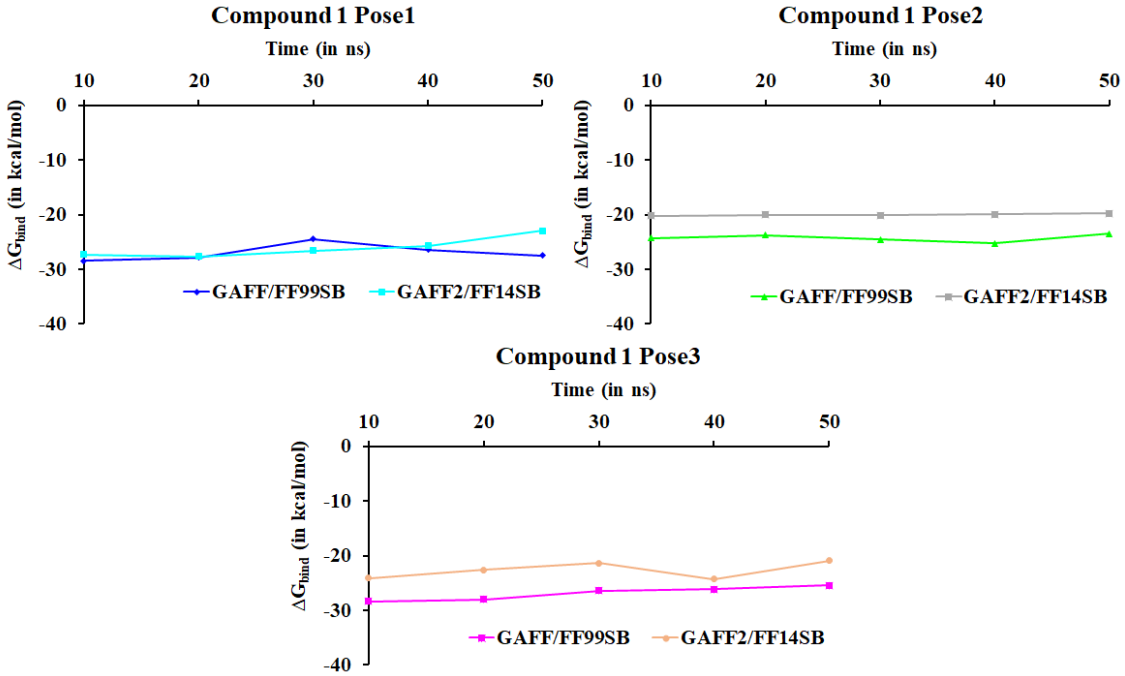

Figure S44. Comparative analysis of the MM/GBSA ΔG_{bind} values calculated over the period of simulation for the three binding conformations (Pose 1, Pose2, and Pose 3) of **compound 1** (bottom row) by applying GAFF/FF99SB,[3,4] and GAFF2/FF14SB.[13,14]

Table S9. Various molecular properties of the compounds selected after virtual screening. (MW = molecular weight; HBA = hydrogen bond acceptor, HBD = hydrogen bond donor, TPSA = total polar surface area, AMR = molar refractivity, nRB = number of rotatable bonds, nAtom = number of atoms, nRigidB = number of rigid bonds).

Compound	MW	LogP	HBA	HBD	TPSA	AMR	nRB	nAtom	nRigidB
1	406.14	0.903	10	2	115.7	106.76	5	48	29
3	405.18	0.053	8	2	103.23	111.18	6	53	27
4	379.15	1.181	7	1	82.44	107.28	6	49	25
8	447.23	1.319	9	3	103.01	113.19	5	62	30
9	383.16	1.392	6	2	71.95	109.29	7	50	24
10	340.14	1.343	6	3	95.5	97.79	8	45	18
12	439.21	0.444	8	3	121.98	113.95	13	63	18
13	467.21	1.162	9	1	113.63	118.06	11	65	23

Table S10. Source library details and CAS registry number (CAS#) for the selected compounds.

Compound	Source Library	ID in the source library	CAS#
1	Enamine	Z70903563	919925-66-9
3	Enamine	Z512963032	1302200-63-0
4	ChemBridge	89402576	1452580-17-4
8	Enamine	Z1310770108	1376138-02-1
9	ChemBridge	85137294	1497566-77-4
10	ChemBridge	9158528	940487-85-4
12	ChemDiv	K938-0746	899704-88-2
13	ChemDiv	L799-1747	1357780-41-6

References

(1) Wang, J. H.; Smolyar, A.; Tan, K.; Liu, J. H.; Kim, M.; Sun, Z. Y. J.; Wagner, G.; Reinherz, E. L. Structure of a Heterophilic Adhesion Complex between the Human CD2 and CD58 (LFA-3) Counterreceptors. *Cell* **1999**, *97*, 791–803.

(2) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham III, T. E.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; et al. Amber 2018. *University of California, San Francisco*. 2018.

(3) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(4) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins Struct. Funct. Genet.* **2006**, *65*, 712–725.

(5) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comput. Chem.* **2002**, *23*, 1623–1641.

(6) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, *105*, 9954–9960.

(7) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(8) Forester, T. R.; Smith, W. SHAKE, Rattle, and Roll: Efficient Constraint Algorithms for Linked Rigid Bodies. *J. Comput. Chem.* **1998**, *19*, 102–111.

(9) Roe, D. R.; Cheatham III, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.

(10) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

(11) Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opin. Drug Discov.* **2015**, *10*, 449–461.

(12) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: An Automated Pipeline for the Setup of Poisson-Boltzmann Electrostatics Calculations. *Nucleic Acids Res.* **2004**, *32*, W665–W667.

(13) Vassetti, D.; Pagliai, M.; Procacci, P. Assessment of GAFF2 and OPLS-AA General Force Fields in Combination with the Water Models TIP3P, SPCE, and OPC3 for the Solvation Free Energy of Druglike Organic Molecules. *J. Chem. Theory Comput.* **2019**, *15*, 1983–1995.

(14) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.